UNIVERSITY OF CALGARY

New Paradigms for Human-Robot Interaction Using Tangible User Interfaces

by

Cheng Guo

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

DECEMBER, 2008

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled “New Paradigms for Human-Robot Interaction Using Tangible User Interfaces” submitted by Cheng Guo in partial fulfillment of the requirements for the degree of Master of Science.

Supervisor, Dr. Ehud Sharlin
Department of Computer Science

Dr. Saul Greenberg
Department of Computer Science

Dr. Gail Kopp
Faculty of Education

Date

Abstract

As technology advances, more and more robots are expected to integrate into our everyday life. However, due to current limitation, we can not communicate with robots directly using languages, gestures and facial expressions. Some type of interfaces is still needed to assist the robots to understand human commands or vice versa. Robots are digitally controlled physical entities that exist in both the virtual realm and the physical world. They are capable of interpreting bits and bytes and converting them into physical outputs to interact with their surroundings. Thus, we believe that an interface that emphasizes physical interaction (physical input mapped to physical output) may be more intuitive to understand and use than the traditional interfaces when interacting with robots.

In this thesis, we propose the use of tangible user interfaces (TUIs) in human robot interaction (HRI) and demonstrate two TUIs that we have created for interacting with a single robot and a team of robots. The first TUI utilizes generic TUIs for controlling the locomotion and posture of an AIBO™ robot dog. The second TUI uses plush toys and allows a single user to interact with multiple robots at the same time. A usability evaluation is conducted for each of these TUIs in comparison with a reference interface that we designed. We believe that one of these baseline systems, a touch-based interface for a team of robots, is a contribution in itself, beyond its value for the TUI comparative study. The consensus from the studies has indicated that TUIs are generally easier to learn, more enjoyable and efficient to use in completing certain HRI tasks.

Publications

Materials, ideas, tables and figures in this thesis have previously appeared in the following publications:

Guo, C., Young, J. E., and Sharlin, E. (2008). **Touch and toys – new techniques for interaction with a remote group of robots**. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI ’09). (To appear)

Guo, C. and Sharlin, E. (2008). **Exploring the use of tangible user interfaces for human-robot interaction: A comparative study**. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI ’08). ACM Press, 121-130. Best Paper Award Nominee.

Guo, C. and Sharlin, E. (2008). **Utilizing physical objects and metaphors for human robot interaction**. In Proceedings of Artificial Intelligence and Simulation of Behavior (AISB ’08). AISB Press.

Acknowledgments

There are many people who have supported and encouraged me throughout this research that I would like to acknowledge.

First and foremost, I would like to thank my supervisor, Dr. Ehud Sharlin: Thank you so much for your trust, guidance, support, encouragement and patience over the past few years. You have provided me the opportunity to work with some of the best researchers in the field of HCI and HRI. I feel incredibly lucky and honoured to be one of your students. You taught me not only how to do research, but more importantly, the correct attitude to face various challenges in life. Thank you!

I want to thank my thesis examiners, Dr. Saul Greenberg and Dr. Gail Kopp for their comments and insightful discussion about my work.

I must acknowledge Dr. Tak Shing Fung for his invaluable help on both of the research works that I have completed and published.

To my friends and colleagues from the Interactions Lab: I am very proud to be a member of the iLab. Thank you all for the friendship and support over the past few years. I want to especially thank the core of the “Ehud’s Gang” members, including James E. Young, Cody Watts and Min Xin. I will never forget the fun times that we had together.

To James E. Young: I am very glad to have the opportunity to work with you together and published our work at CHI. (どうもありがとうございます!)

To Xi Chen: Thank you for your encouragement and support (including the distractions). Thank you very much for always been there for me.

To my mother, Shu Fang Ou, and my father, Ke Li Guo: I am extremely grateful for the opportunities, supports and love that you have provided. Without your care and support, I would not be able to complete both of my Bachelor’s and Master’s degree in Canada.

Finally, I would like to acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) and the University of Calgary for their financial support.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Robots are digitally-controlled physical entities that exist in both the virtual realm and the physical world. They are capable of interpreting bits and bytes and converting them into physical output. They are also capable of sampling and sensing physical phenomena and translating them into digital information. With the advance of technology, new functionalities have been added to modern robots, enhancing their abilities to interact with a wide range of physical objects and granting them the ability to communicate with humans using human language. Over time, a split view on the role of robots has developed. One view treats robots as tools (Yanco and Drury, 2004). Proponents of this view see robots as puppets of their human controllers, designed only to accept and execute commands. The opposing view sees robots as companions which can help to fulfil a human partner's social needs. Researchers who hold this view have been analysing education, health care, and the home environment in order to develop robots that are capable of social interaction. (Brea zeal, 1999; Fiorini et al., 1997)

Unlike some advanced sociable robots which are capable of interpreting visual and audio commands, many robots can only communicate with humans through blinking LEDs, or console output. In order to interact with these robots, mice, keyboards and graphical user interfaces (GUI) are commonly used. While these interfaces may be useful for performing many tasks on a PC, their generic affordances make them a poor choice for interacting with robots. Therefore, in this thesis, we introduce our vision of using tangible user interfaces (TUI) for interacting with non-autonomous robots (Yanco and Drury, 2004). We also describe the design, implementation and evaluation of two novel TUIs-based robotic interfaces: the Wiibot and Toy interfaces which we have created for performing a set of robotic tasks. These two interfaces have been evaluated through extensive design and performance testing against non-TUI reference interfaces. These reference systems, and the related comparative experimentations, are also described in this thesis.

In this first chapter, we briefly introduce the background and motivation behind our research. We describe the challenges that users face when interacting with robots via traditional user interfaces. Then, we define our research goal and explain our approach to addressing it. Finally, we conclude this chapter with an organizational overview of the remainder of this thesis.

1.1 Background

Humans possess the rare ability to create tools to help them to achieve their goals. Even in the early Stone Age, humans were able to utilize physical matter as tools for lighting fires and hunting animals. With the discovery of new materials, such as bronze and iron, humans were able to create a variety of new artifacts for various tasks. By combining tools with sources of energy, such as steam and electricity, people have been able to achieve certain level of automation so that their tools can work “by themselves”.

Looking back at human history, it is easy to see how inventors have not only created new tools but also refined existing tools to achieve higher levels of usability, efficiency, and aesthetic value. For instance, the sundial – the world’s first timekeeping device – is capable of measuring individual hours in a day. The earliest known sundial was build by ancient Egyptians around 3500 BC (Encyclopedia Britannica, 2008). Over thousands of years, people have refined the process of clock-making and perfected the technique. Today, clocks have shrunk from giant obelisks to wearable wrist pieces – some of which are accurate to within nanoseconds. Nowadays, certain brands of wristwatches are recognized as luxury goods and a sign of social status. Depending on their intended purpose, clocks have been modified into many shapes and sizes. For example, in short running races, referees may use stopwatches to record the athlete’s finishing times. But large-screen digital stopwatches are used to display finishing times to broader audiences. This shows how the same kinds of tools may possess different physical properties, visual presentations and affective attributes depending on their usage.

Since the invention of the personal computer, many different kinds of tools have been invented for interacting with computers effectively and intuitively. Eventually, the keyboard, screen and mouse became widely adapted as the primary, generic interfaces for PCs. Although these tools are suitable for many kinds of applications, their physical shapes and affordances (Norman, 1997) can impose difficulties on users when used in tasks for which they are not well-suited. For instance, drawing a picture on the PC using a mouse is very cumbersome. Thus, digital pens and tablets have been developed to allow precise drawing or quick jotting in a more traditional and intuitive way. Following the evolutionary path of human-made tools, one would expect to see a growing variety of dedicated physical tools being created to facilitate interactions between humans and robots.

Human-robot interaction (HRI) is a relatively new sub-area within the field of human-computer interaction (HCI). (Kiesler et al., 2004) The tools for interacting with robots are still naïve and their abilities are limited. The most commonly used interfaces for interacting with robots are still the joystick, mouse, keyboard and graphical user interface (GUI). The combination of these interfaces has been treated as a “universal controller” for interacting with robotic applications by many HRI researchers. However, as robots evolve, it is likely that the tools we use for controlling and interacting with them will evolve as well. Similar to our clock and digital pen examples, it is inevitable that people will either refine their current tools or invent new ones to enhance the efficiency and reduce the difficulty of human-robot interactions. In the near future, we should expect to see a set of more dedicated tools that are designed specifically for working with a particular type of robot or for a particular kind of robotic task.

1.2 Motivation

Interacting with a robot is dramatically different from using a personal computer to produce text documents, send emails, browse the web and perform the many other tasks in which the mouse-keyboard-screen interaction approach excels. First of all, robots are physical entities that have the ability to sense and to alter the physical environment around them. Depending on the size and weight of the robot in question, it may have the potential to damage objects or people around them. Thus, it is crucial for the user to be aware of the physical state of the robot and its surroundings (Drury et al., 2003). This spatial awareness problem is exacerbated when multiple robots need to be controlled by a single operator. The second major problem which inhibits smooth human-robot interaction is a lack of intuitive communication. When people share space and work together, information is exchanged using speech, facial expressions, eye gaze, touch and body language. Like humans, robots are spatially-active physical entities. However, robots understand none of these communication methods that are commonly used by humans. The current technology is not advanced enough to permit information exchange between humans and many robots without some intermediary device. Therefore, the tools that we use to bridge this communication gap directly affect the quality of the human-robot interaction.

To date, some effort has been made to explore the possibilities of using speech, gestures and other common human-to-human interaction techniques for interacting with robots. (Hasanuzzaman et al., 2004; Rogalla et al., 2002) However, such efforts have been primarily focused on the technological aspects which support these interaction techniques, rather than the interaction itself. Consequently, little work has been done to evaluate these “natural” interfaces against the traditional input devices for a specific set of robotic tasks. Therefore, it is still difficult for designers to categorize which interfaces may be better to use given a specific set of robotic tasks and under a specific set of circumstances. One exception is the work been done by Quigley et al. in 2004. They compared seven different types of interfaces for controlling an unmanned aerial vehicle (UAV). Their results show that interfaces that are “tailored to the physical and cognitive requirements of the application” outperform the traditional graphical user interface.

To explore alternative user interfaces for intuitive robot control, we introduce our vision on utilizing tangible user interfaces (TUI) for specific HRI tasks. The expression “tangible user interface” was firstly popularized by Ishii and Ullmer (1997). Since then, TUIs have gained much attention in the Human-Computer Interaction (HCI) community. TUIs take advantage of our innate abilities in manipulating objects for controlling virtual entities. By touching and moving physical objects which embody digital content and function (Dourish, 1997), users are able to interact with digital objects, information and actions directly and intuitively.

The benefits of using TUIs for HRI are manifold. For instance, the shape, size and weight along with other physical properties of an object imply the way we interact with it. If we can appropriately map the physical properties (such as physical constraints) of a robot to the physical properties of a TUI, then the potential functionalities and mechanism of a robot can be directly revealed to the operator. Moreover, the spatial orientation and the position of a physical object in relation to its surroundings can expose additional information and provide interaction insight and task awareness to the manipulator. Therefore, by taking the advantage of the physical properties of TUIs, we may offer additional awareness information and other activities that traditional user interfaces do not afford for HRI tasks.

1.3 Research Questions

In this thesis, we explore the use of tangible user interfaces for human robot interaction. To be more specific, we investigate the use of physical objects as functional and content "handles" for interaction with robots. My research will not only focus on one-to-one cases (i.e. one human operator interacts with a single robot) but also on one-to-many (i.e. one human operator controls multiple robots) scenarios. In order to investigate how the novel TUI-based interfaces we developed fare compared to non-TUI interfaces we had to design a baseline system, including a new touch-based human-robot interface. In this thesis we address the following research questions:

1. ***Why apply tangible user interfaces to HRI?*** What kind of benefit do tangible user interfaces offer for tackling HRI-specific challenges?

2. ***How can we apply tangible user interfaces to allow a single operator to interact with a single robot?***
3. ***How can we apply tangible user interfaces to allow a single operator to interact with a team of robots?***
4. ***What are the advantages (or disadvantages) of using tangible user interfaces for completing HRI tasks compared to more orthodox, and to non-tangible user interfaces?***

1.4 Approach

To answer the research questions stated above, we present the design, implementation and evaluation of two new TUI-based robotic interfaces which we created.

1. *TUIs for HRI - one operator interacts with one robot*. To test the feasibility of applying TUIs to HRI, we used a pair of Nintendo Wii™ Controllers as TUIs for controlling a Sony AIBO™ robot dog in a set of tasks. This system is built for demonstrating the use of TUIs for HRI on a real robot. We also use this scenario to determine the advantages and disadvantages of using generic TUIs for robotic tasks.
2. *TUIs for HRI - single operator interacts with multiple robots.* Building upon the success of the single robot control system, we extend our original idea to a TUI-based interface for controlling a group of heterogeneous robots. We used a high-resolution tabletop computer and a set of children's toys for the TUI-based interface, affording an intuitive mapping between the user's movement and the robots' reactions, and providing visual HRI awareness of the robots' states to the user.
3. *Comparative user studies*. We conducted two user studies to test the usability of both our TUI systems. Two different reference systems have been designed and implemented for the single and multiple robot case for comparison with their TUI counterparts: a keypad-based system for the single robot case, and a unique touch-based system for the robotic group case. More than fifty participants have been recruited and participated in these user studies, and the experimental results have been analysed via quantitative as well as qualitative methods.

1.5 Contributions

Using the proposed approach for answering the aforementioned research questions, this thesis makes the following five contributions:

1. Proposing the use of TUIs for HRI. As far as we know, we are the first ones who suggest merging the area of TUIs and the area of HRI.
2. Demonstrating the use of TUIs for interacting with a single robot. As far as we know, it is among the very first published works of TUIs for interacting with a real robot.
3. Demonstrating the use of TUIs for interacting with a team of heterogeneous robots. As far as we know, it is the first published use of TUIs for interacting with a group of robots.
4. Demonstrates the use of a touch interface for interacting with a team of robots. As far as we know, it is the first research work that utilized a touch-enabled tabletop computer for interacting with a group of heterogeneous robots.
5. Comparing the TUIs we created for single robot interaction with a traditional keypad-based interface in terms of usability. Through the user study we conducted, we have collected data from twenty eight participants to reveal the advantage of using TUIs for HRI tasks.
6. Comparing the TUIs we designed for multiple-robot interaction against a touch-based interface. By conducting a user study, data was collected from participants and used this data to create a list of implications for designing tabletop-, touch- and TUI interaction techniques with a remote group of robots.

1.6 Thesis Overview

The remaining of this thesis describes in detail the research outlined above. In Chapter 2, we provide a brief review of related HRI efforts. In Chapter 3, we outline our vision of applying TUIs to HRI and detail our motivation behind using TUIs for robotic interfaces. In

Chapter 4, we describe the design, implementation and evaluation on the first TUI system that we created for interacting with a single robot using the Nintendo Wii™ controllers. In Chapter 5, we describe the architecture and implementation details, as well as a preliminary study, on the usability of the second TUI system we created for interacting with multiple heterogeneous robots using toys. Finally, in Chapter 6, we discuss the implications of the human-robot interfaces we designed and evaluated on future efforts in this domain, and conclude this thesis with a discussion of possible short and long-term future efforts towards the research goals we outlined earlier.

Chapter 2. Related Work

In this chapter, we present a brief overview of the common interfaces that are used for human-robot interaction (HRI). Next, we discuss tangible user interfaces (TUIs) within the context of human-computer interaction (HCI). Then, we review the particular challenges of remote robot control and multiple robot control. These topics serve as basis for the discussion of applying TUIs to HRI control tasks, as presented in the next chapter.

2.1 Human-Robot Interfaces

Currently, there are a variety of robots available on the market. Examples include robotic arms that are used in space station assemblies to explosive ordnance disposal robots dispatched onto battlefields. If we can draw an axis to categorize them based on their usage, on one end of the axis, we have robots that work as tools (i.e. iRobot PackBot, 2008.) to help humans to accomplish certain tasks. On the other end of the axis, we have robots that act as our companions for entertainment (i.e. Kozima et al., 2007) and social interaction (i.e. Nabaztag, 2008) purposes.

Yanco and Drury (2004) categorized robots based on their autonomy level. It represents a similar view as the aforementioned axis analogy. According to their definition, a robot’s autonomy level can be measured as the percentage of task time in which the robot is carrying out its task on its own. In correspondence, the amount of intervention required for a robot to function is measured as the percentage of task time in which a human operator must be controlling the robot. These two measures, autonomy and intervention, sum up to 100%. In our research, we focus on exploring intuitive control methods for manipulating robots which have 0% autonomy and 100% intervention ratios.

Even with today’s technology, most robots cannot interpret our language, facial expression and gestures. Interfaces are still needed to translate human commands to robot actions. To interact with robots, mice, keyboards and graphical user interfaces have been “inherited” from the field of HCI to the field of HRI for this purpose. However, robots’ capability of moving in the real world and interacting with physical objects make them a fundamentally different entity from traditional computing devices (Breazeal, 2003; Kiesler and Hinds, 2004; Norman, 2004; Young et al., 2008). Therefore, traditional human-computer interfaces may not satisfy our needs at handling HRI tasks. To bridge the communication gap and allow for a more intuitive and natural interaction experience, HRI researchers have put much effort into creating new types of interfaces. The following sections detail some of the common interfaces used in the field of HRI for interacting with robots.

2.1.1 Naïve Interfaces

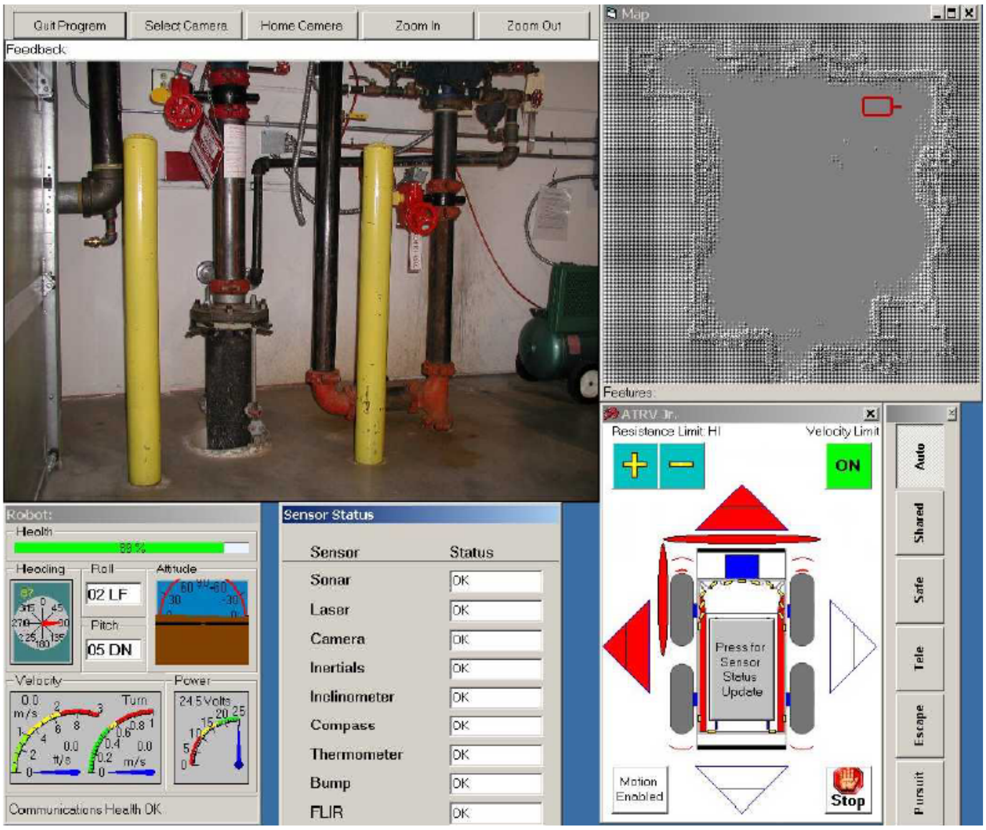
Robots that are designed to work autonomously or with less emphasis on their communication capability are often equipped with simple indicators and buttons. These robots usually use a combination of visual and audio awareness indicators such as light-emitting diodes (LEDs) and loudspeakers to notify the human operator about their current state and malfunctions. For example, Roomba (iRobot, 2008) is an autonomous vacuum cleaning robot, which works in domestic environments. When mechanical problems occur, it uses different color combinations of flashing light and beeps to indicate error codes to its owner. To communicate with the Roomba (e.g., issuing a command to make the Roomba stop), the human operator needs to press a button or a combination of buttons to trigger a particular function.

On one hand, the robot operator is easily notified through the visual and audio feedback when the robots demand attention (assuming the robot and its operator are co-located). However, on the other hand, without a comprehensive user manual to explain the meaning of each light and beep combination and error codes, it is impossible to tell what kind of information the robots are trying to convey. The lack of expressiveness and intuitiveness make naïve interfaces hard to interact with from the users' perspective. Therefore, these interfaces are not suitable for interacting with robots that require high intervention ratio (Yanco and Drury, 2004).

2.1.2 Standard HCI Interfaces for HRI

Standard HCI interfaces, such as graphical user interfaces, mouse, keyboard, and joysticks, are commonly used in the field of HRI. For example, Figure 2.2 demonstrates a GUI for controlling a search and rescue robot. The on-screen interface consists of a streaming video window with buttons (top-left), a map (top-right), a robot status panel (bottom-left), a sensor status panel (bottom-mid) and a control panel (bottom-right) where the operator can navigate the robot. In combination with a mouse and keyboard, the operator is able to interact with robots in a way that is similar to interacting with a generic desktop application. In special cases, knobs, gamepads and joysticks are used instead of a mouse and keyboard. For example, Figure 2.1 shows a PackBot EOD's (Explosive Ordnance Disposal) portable command console. The knobs on the panel allow the operator to navigate the robot and adjust the height of the robotic neck. The buttons on the unit allow for adjusting screen display, speaker and microphone volume and other functionalities. People also use joysticks for navigating robots. For instance, joysticks are commonly used in controlling unmanned air vehicles (UAVs). (William, 2004; Quigley et al, 2004) We have also seen examples of using joysticks for navigating custom build robots, such as the GestureMan

Figure 2.1 – Controlling robots using knobs and gamepad (from www.irobot.com)



Sensor	Status
Sonar	OK
Laser	OK
Camera	OK
Inertials	OK
Inclinometer	OK
Compass	OK
Thermometer	OK
Bump	OK
FLIR	OK

Figure 2.2 – A Graphical User Interface for Remote Robot Control (from Yanco et al., 2004)

robot. (Kuzuoka et al, 2004)

There are some benefits of using standard HCI interfaces for HRI tasks. First, the ease of deployment coupled with a rich variety of visualization tools makes GUIs a popular choice over many other types of interaction techniques. Different visualization techniques can be used to effectively display sensor data (Murta et al., 2000; Ricks et al., 2004). Also, on-screen widgets can change their states dynamically to reflect the current status of the robot. For example, if a robotic function is unavailable due to mechanical problem, the one-screen button for triggering this function can be dynamically disabled to prevent confusion. Secondly, the generic nature of the GUI, mouse and keyboard paradigm (from here on we will refer to this paradigm as the traditional user interface) potentially makes it universally available to all robotic platforms and tasks. Thirdly, two-dimensional robot tasks can be easily mapped to traditional UI. For example, many researchers use the traditional UI for two-dimensional navigation tasks. The “W, A, S, D” keyboard mapping or mouse clicking on a top-down map view are straightforward methods for controlling robots that are capable of moving in a two-dimensional space.

Although it is natural to carry over the traditional user interface (traditional UI) from HCI and use it for HRI tasks, the intuitiveness and effectiveness of the traditional UI in controlling robots remains debatable for three primary reasons. Firstly, the mouse, keyboard, and graphical user interfaces separate user input from computer output, uncoupling action and perception space, and potentially breaking the flow of users’ cognitive engagement when performing certain tasks. (Faisal et al., 2005) For instance, when typing on a keyboard, most people need to look at both the keyboard and the computer screen to ensure they entered the correct character. In terms of telerobotics, human operators have to solely rely on the image and sensor data transmitted back by the robot to determine their next operation. Constantly switching attention back and forth between the input device and the data display screen is not ideal, especially when the robot is in critical conditions. Secondly, the motor skills required for manipulating a mouse and typing on a keyboard are not intuitive to learn. A sufficient amount of time is required for people to memorize the layout of the keyboard and repeatedly practice in order to type without looking at the keys. When it comes to robot control, the longer it takes a human operator to master certain motor skills, the greater the cost (time, money and labor) of training will be. In addition, the amount of attention the operator needs to spend on the input device is likely to be higher, which may hinder the overall performance. Thirdly, two-dimensional traditional UIs limit people’s spatial abilities when interacting with three dimensional objects. Thus, use of the traditional UI can be difficult to control a robot that is capable of moving in three-dimensions, such as, controlling an unmanned aerial vehicle (UAV) using the traditional UI. (Quigely et al., 2004; Drury et al., 2006)

2.1.3 Gesture-Based Interfaces

Gesture-based interfaces are one of the most intuitive and unobtrusive method that allows people to communicate with robots. The idea behind this type of interface is to map a set of human hand and body gestures to a set of robot actions. To enable a computer to track human hand and body gestures, computer vision and glove-based input devices are commonly used for this purpose. “One-to-one” and “one-to-many” mapping are usually used between human gestures and robot actions. For example, for one-to-one mappings, a robot would simply mimic its operator’s movement (Uskarci et al., 2003) or accomplish simple commands, such as stand up (Hasanuzzaman et al., 2004). For “one-to-many” mappings, the human operator would perform a simple gesture (such as pointing to an object), then the robot would firstly locate the object of interest, and then it will reach the object and eventually pick it up (Becker et al., 1999).

In the field of computer vision, many algorithms have been developed that track human hand and body gestures to allow for robots interact with humans in a natural and unencumbered fashion (Kortenkamp et al., 1996; Waldherr et al., 2000; Hasanuzzaman et al., 2004). The human operator’s gestures are captured by either one or more cameras. Depending on the computation power of the robotic platform, the cameras can be either placed on the robot itself (if the robot is capable of performing real-time image processing and co-located with the operator) or in a controlled environment (an area surrounded by fixed cameras). If the cameras are placed in a controlled environment, then a computer is commonly used as mediator for interpreting the captured gestures and controls the robot to act accordingly. By using different segmentation techniques, researchers can extract the body parts of interest from the image background. Then, the extracted image is analysed by pattern-matching algorithms to determine which gesture the human operator is performing (Theis et al., 2001; Hasanuzzaman et al., 2004).

With computer vision based systems, robot operators do not need to wear any instruments for monitoring their gestures. Thus, robot operators are freed from learning and mastering any tools when interacting with robots. By using hand and body gestures, an operator can command a robot directly. However, computer vision based systems have their own disadvantages. In a controlled environment, cameras need to be setup and calibrated before use. Depending on the computer vision algorithm implemented, vision-based interfaces can be sensitive to the changes in background colors and intensity of ambient lights, which may greatly affect detection results. Also, the human operator needs to remain close to the cameras in order to obtain good detection result from the system.

Another tool commonly used to capture hand gestures are glove-based input devices. These devices usually require human operators to wear a glove with sensors on them which monitoring their finger joints and hand's position and orientation. A computer is usually connected with the glove for receiving and analysing the data transferred from the glove. By analysing the angles between finger joints and the motion of the hand, gesture recognition algorithms are able to determine which gesture the human operator is performing. Then, the corresponding commands are sent to the robots. These interfaces are usually used to control either real dexterous robot hands or virtual robots (Sturman and Zeltzer, 1994). For instance, one well-known example of using a glove-based interface for controlling real robot hands is the Robonaut project (Bluethmann et al., 2003) (Figure 2.3). Bluethmann et al. (2003) adopted a master-slave system approach which requires the human operator to wear gloves equipped with Polhemus trackers for detecting arm and hand positions. The Robonaut operator remotely controls the Robonaut from a distance without physically touching it.

Figure 2.3 - Robonaut's teleoperation hardware (from Bluethmann et al., 2003)

Similar to vision tracking methods, glove-based tracking methods also have technical limitations. Depending on the size of the hand, some glove-based interfaces require precalibration for each user. Similar to vision-based interfaces, the human operator can use static finger postures and hand motions to control robots directly. But some gloves impede finger movements due to the material they are made of and the design of the glove (Sturman and Zeltzer, 1994).

Besides the aforementioned technical challenges for gesture-based interfaces (including both vision and glove-based methods), another limitation of these interfaces come from the human body. Our gestures and movements are limited by our skeletal structure. For example, we cannot turn our head 360 degrees. However, robots can perform gestures that we are not capable of. For instance, some robot hands' wrists can rotate 360 degrees like a drill. It is impossible to map our hands' wrists movements to this kind of action. In addition, muscle fatigue would prevent us from performing repetitive and prolonged activities. Therefore, we need to avoid mapping repetitive and prolonged gestures to robot actions when using gesture input systems. For example, if we were controlling a drilling robot arm to drill a spot on a wall using one of our arms, we need to keep our arm fixed until the drilling operation is done. This is not possible if we have to drill for a lengthy period of time.

2.1.4 Exoskeleton Interfaces

Exoskeleton interfaces is a "mechatronic system designed around the shape and function of a human body, with segments and joints corresponding to those of the person it is externally coupled with" (Wikipedia, 2008). Exoskeleton interfaces are usually worn by a human operator to either augment their physically strength in labor works or teleoperate a robot at a remote place. One early example of the exoskeleton system is the extender system demonstrated by Kazerooni (Kazerooni, 1990). In Kazerooni's research, the human operator wore a robotic arm to directly apply *mechanical power and information signals* (Kazerooni, 1990) to the robot. By measuring the dynamic contact force applied by the human operator, the robotic limbs are able to amplify that force for performing heavy duty tasks that normal human strength would not be capable of. A more recent project that utilized the idea of the exoskeleton interface is the BLEEX project (Berkeley Robotics Laboratory, 2008). BLEEX is a pair of robot leg braces and a backpack frame that can be attached to the human body. It provides extra support for people who need to carry weights that are beyond their capacity. As the example shown in Figure 2.4, the person is able to walk in a natural fashion with a 70-pound bag on his back (Berkeley Robotics Laboratory, 2008).

Figure 2.4 – Berkeley Lower Extremity Exoskeleton
(from http://bleex.me.berkeley.edu/bleex.htm)

2.2 Tangible User Interfaces

Tangible user interfaces (TUIs) exploit embodied interaction (Dourish, 2001), coupling physical objects with computerized qualities, and ideally empowering users with simple and natural physical interaction metaphors.

The notion of tangible user interfaces (Ishii and Ullmer, 1997) is based on Fitzmaurice et al.'s earlier Graspable User Interfaces effort (Fitzmaurice et al., 1995). Fitzmaurice and Buxton (1997) have conducted an experiment which allowed users to use "Bricks" as physical handles to direct manipulate virtual objects. Their study has shown that "a space-multiplex input scheme with specialized devices can outperform a timemultiplex (e.g., mouse-based) input design for certain situations”. (Fitzmaurice and Buxton, 1997) Later, Ishii and Ullmer (1997) proposed the term Tangible User Interfaces and addressed the importance of both the foreground interaction which consists of using physical objects to manipulate virtual entities and the background interaction which happens at the periphery to enhance users’ awareness using ambient media in an augmented space.

Ishii and Ullmer (1997) defined TUI as “seamless coupling everyday graspable objects with the digital information that pertains to them”. In their research (Ishii and Ullmer, 1997), they demonstrated the use of physical objects for controlling virtual entities. For example, users are able to alter the scale of a digital map directly by either sliding two cylinders close together (to zoom in) or separate them apart from each other (to zoom out) (Figure 2.5).

Figure 2.5 – Scaling device for zoom in/out on a digital map (Ishii and Ullmer, 1997)

Successful TUIs take full advantage of physical objects’ properties to aid the users in the interaction process. One of the most important properties that physical objects provide is their affordances (Norman, 1988). According to Norman (1988), affordance is defined as “the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possible be used”. It suggests the usage of the objects in relation to its actor’s physical capabilities through its shape, size and weight along with other physical properties (such as color and tactile feeling). Research (Gibson, 1969; Adolph et al., 1993) has shown that “very young infants are able to perceive the affordances provided by the physical layout of surfaces in their environment, including those that support locomotion, those that afford falling, and those that afford collision”. Moreover, by 5½ months of age, infants are able to perceive the affordances for action of everyday objects. They can discriminate between the correct and incorrect use of common objects in the context of everyday actions (Anasagasti et al., 2002). Thus, we can take the advantage of our innate skills at observing and learning how to interact with physical objects in TUI design, which may reduce the number of new motor skills we need to acquire and lower the number of mistakes we make during the interaction.

Another important property that successful TUIs exploit is spatiality. Both humans and TUIs are spatial beings. They exist in the physical world, occupy physical space and can move or be moved around within the physical environment. Our innate skills of spatiality allow us to perceive spatial qualities, such as the distance between an object and ourselves (Hornecker and Buur, 2006). This information may aid us in the process of interacting with other objects that are located in the same environment. To achieve a good spatial mapping between a physical objects and a virtual entity, a good degree of integration and compatibility (Beaudouin-Lafon, 2000) is also needed (Sharlin et al., 2004). Beaudouin-Lafon defines the degree of integration as “the ratio between the degree of freedom (DOF) provided by the logical part of the instrument and the DOFs captured by the input device”. For example, using a 2D device, such as a mouse, to rotate a 3D object in Maya (Autodesk, 2008) has a degree of integration of 3/2. He also defines the degree of compatibility as “the similarity between the physical actions of the users on the instrument and the response of the object”. For instance, dragging an icon on a computer desktop has a high degree of compatibility since the icon follows the movement of the mouse. A good physical representation and a spatially congruent mapping can afford an appropriate degree of integration and high degree of compatibility, which in turn, make TUIs’ functionality easy to perceive and to operate by users. An example of a TUI system that demonstrates a good spatial mapping between the TUI object and its use in the digital application is the tangible avatar project (Yamashita et al., 2007). Yamashita et al. used a physical doll as an avatar to represent its controller in a 3D virtual world. (Figure 2.6) By moving and changing the orientation of the physical avatar on a table surface, users are able to explore the 3D virtual world easily using their spatial perception.

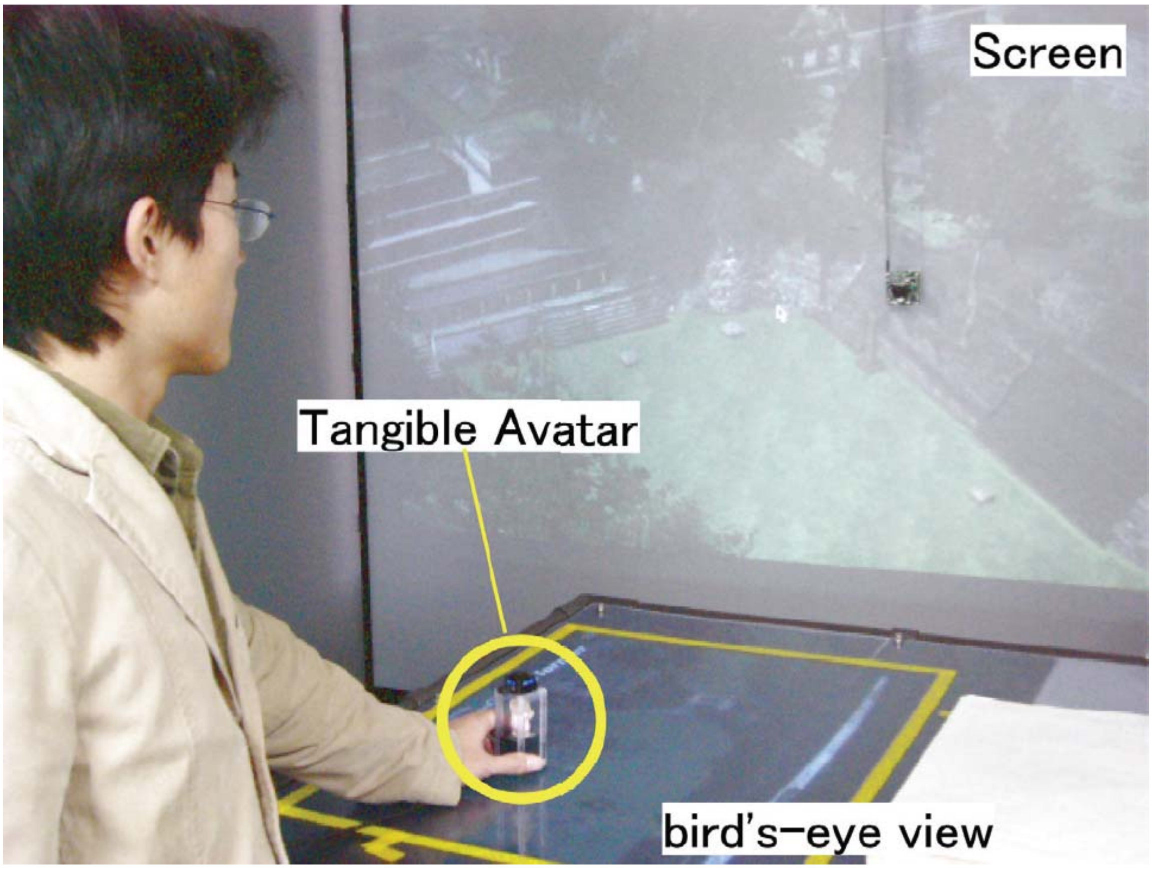


Figure 2.6 Using the tangible avatar to explore a virtual world. (Yamashita et al., 2007)

Thirdly, TUIs potentially allow for unification of action space (input) and perception space (output) (Sharlin et al., 2004). Unlike traditional UIs which separate the action and perception space (such as typing on a keyboard and then looking on a screen for confirmation), the project shown in Figure 2.5 demonstrates a good example of unification of user input and system output. For instance, as a user moves the cylinders close together, the virtual map zooms in on the spot where the tool is placed in real time. This system provides immediate feedback to the user when he/she interacts with the virtual map using different tools.

Fourthly, successful TUIs support “trial-and-error” activity (Sharlin et al., 2004). For instance, many text editor programs allow for undo operation. However, users can not revert the program back to the state five operations earlier without undo the four previous operations. TUIs can provide multiple access points to the system, which in turn, supports “trail-and-error” activity. For example, the Senseboard project (Jacob et al., 2002) allows users to place magnetized pucks (which represent events) on a scheduling board to alter events that are recorded in a scheduling application (Figure 2.7). Since each event puck can be directly accessed physically, the removal of a previously added event can be accomplished by taking a puck off the Senseboard in one step in compare to a series of linear undo operations.

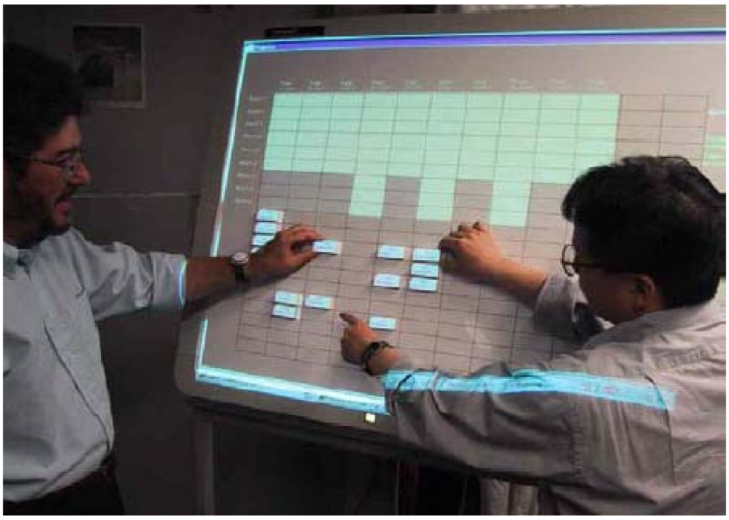


Figure 2.7 – Senseboard (Jacob et al, 2002)

To conceptualize and better understand the idea of tangible user interface and tangible interaction in general, two pieces of works need to be mentioned here for their contributions. In 2004, Fishkin presented a spectrum-based taxonomy of tangible interfaces. He uses the levels of embodiment and the levels of metaphor for categorizing TUI applications. He defines the levels of metaphor as the levels of likeness of system effect of a user action analogous to the real-world effect of similar actions. He argues that there are two types of metaphor that TUIs afford: the “metaphor of noun” (i.e. TUIs which appeal to the shape of an object) and the “metaphor of verb” (i.e. TUIs which appeal to the motion of an object) (Fishkin, 2004). Based on this categorization, he separates the levels of metaphor into five different scales, None, Noun, Verb, Noun and Verb, and Full (i.e. the virtual system is the physical system). Researchers can use his scales to categorize their work and use it to compare and contrast with other disparate research efforts in a systematic way.

Hornecker and Buur (2006) produced a framework for guiding researchers to evaluate different TUI designs and to understand the user experience and social aspects of different TUIs. They divided their framework into four themes: tangible manipulation, spatial interaction, embodied facilitation and expressive representation. Researchers can use these themes to evaluate TUI designs in terms of system input/output, interaction style, visual and physical representation and social interaction.

2.3 HRI in Telerobotics

Although robots are becoming more intelligent and capable of handling simple tasks (e.g. vacuum cleaning) autonomously, they still need humans' assistance and intervention when dealing with complex situations. So far, artificial intelligence is not advanced enough to replace human cognitive capabilities for task planning and human sensorimotor capabilities for control (Durlach and Mavor, 1995). Therefore, a human-machine interface that allows the inclusion of human operators as part of the robot control system is needed. In this section, we focus on discussing the common issues faced in teleoperating a single and a team of robots that have 0% of autonomy and 100% intervention ratio (Yanco and Drury, 2004).

One persistent problem that interface designers need to face is how to let human operators maintain HRI awareness through human-machine interface. Drury et al. (2003) defined HRI awareness as "the understanding that the human has of the location, activities, status, and surroundings of the robot; and the knowledge that the robot has of human's commands necessary to direct its activities and the constraints under which must operate." When teleoperating robots, operators have to solely rely on the sensor data and video image provided by the robots to perceive the remote environment. Our senses are completely replaced by numbers and images the robot perceives. Therefore, how to maximize human sensory through user interfaces becomes a challenging design issue that HRI researchers have to tackle.

Several researchers have suggested ways to improve user interface design for remote robot control. Goodrich and Olsen’s (2003) work provides a general guide on how to design an effective user interface for HRI. Yanco et al. (2004) summarized their findings from a major robotic competition, and complied a guideline for improving operator’s situation awareness in HRI interface design. To broaden the view of HRI researchers in interface design, Richer and Drury (2006) had summarized and formed a video game-based framework that can be used to characterize and to analyse robotic interfaces.

Some robotic applications require one human operator to control a team of robots (Ferketic et al., 2006). In this case, to provide enough situation awareness of an entire robot team becomes an even more challenging issue compare to the single robot control scenario. To address this problem, Kaminka and Elmaliach (2006) suggested a GUI interface which they call "relation tool" for visualizing the relative position of each robot within a tightly-coordinated robot team (Figure 2.8). Jones and Snyder (2001) used the real-time strategy (RTS) game interface paradigm as a design philosophy for supervising and controlling multiple complex robots in a simulated environment.

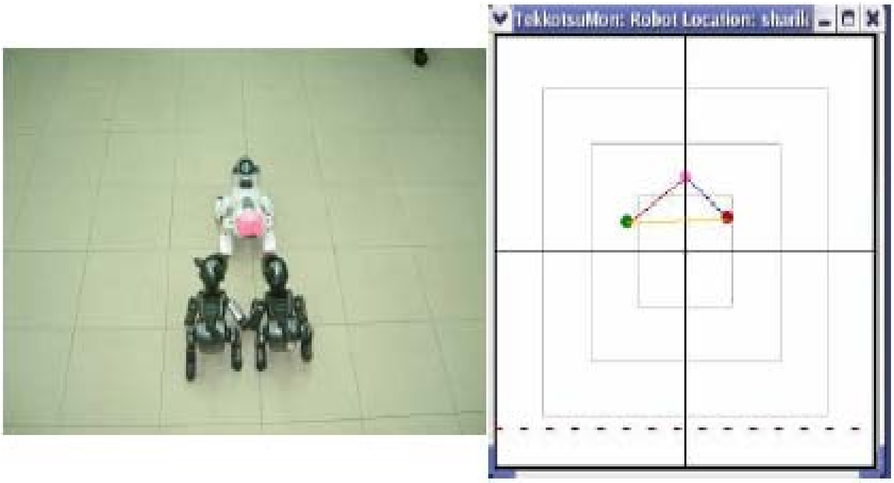


Figure 2.8 – Ground Truth (left) Relation Tool for visualizing the team formation (Right) (Kaminka and Elmaliach, 2006)

Another problem that is specific to the area of multiple robots control is the prolonged time needed to solve *call-requests*. A call-request is initiated by a robot when it require assistant from operators (Kaminka and Elmaliach, 2006). When a team of robots work collaborately on a task, one single call-request may trigger a series of call-requests. It is up to the human operator to prioritize the call-requests and deal with them linearly. Therefore, how to minimize the duration of each call-request within a team of robots is a challenging issue. Kaminka and Elmaliach (2006) proposed two novel distributed methods to tackle this problem. In their research, they asked a participant to control a team of robots to move from one location to another in a pre-defined formation. If a robot is left behind during the transportation, the participant needs to locate and "rescue" that robot. If a "searching for missing robot" call-request is initiated, instead of waiting for the human operator to come up with a plan, the other robot teammates would use their knowledge of where the missing robot was previously located in the group formation to look for it autonomously.

When controlling a group of robots, group-level operations such as, travelling together or gathering, are likely to be performed. A lot of research has been done in the area of artificial intelligence on how to make autonomous robots to form a particular shape based on algorithms (Cao et al., 1995). However, not much research has been done on how to allow human operators to control robot group behaviours manually.

2.4 Summary

In this chapter, we have described the background behind our research. We have reviewed different kinds of HRI interfaces and identified their strength and weakness in interacting with robots. Then, we briefly reviewed the history of TUIs and explained the benefit of utilizing TUIs for HCI tasks. After that, we summarized the problems that are commonly encountered in the field of telerobotics and the current works that try to solve these problems. Although the topics reviewed in this chapter may not seem to share much common ground, these literatures serve as a basis for us to explain our vision of applying TUIs to HRI in the next chapter. In Chapter 4 and 5, we present two TUI prototypes that we created for interacting with a single and a group of robots along with evaluations to further demonstrate our vision of combining TUIs and HRI together.

Chapter 3. Design Approach: Applying TUIs to HRI

3.1 TUIs for HRI: Design Motivation

To address some of the HRI challenges described in the previous chapters, we need to search for alternative HRI interfaces rather than unquestioningly adopting the orthodox mouse, keyboard and GUI paradigm so common in HCI. We see a great potential in using TUIs for supporting natural HRI for the following reasons:

1. *Physical Input to Physical Output.* Since TUIs are physical entities, a physical action applied to them by the user can be directly associated with a robot action. For instance, if we can use a deformable object to control a deformable robot, then the actions we perform on the deformable object (e.g. squeeze, stretch, etc…) can directly affect the shape of the deformable robot. In this case, the consequences of the actions preformed by users can be easily understood and predicted because their immediate physical actions are mirrored by physical robot actions.
2. *Physical feedback.* TUIs are capable of conveying some physical feedback that traditional UI cannot, and are expected to provide even higher level of physical output expressions in the future. Vibrators and other electrical device can be incorporated within TUIs to provide tactile feedback directly to their holders. For instance, when users interact with remote robot, guiding it to push an object, with the traditional UI, the user has to rely on video feedback and sensor data to determine whether the robot has touched the object or not. However, with a TUI approach, the TUI itself can physical react (e.g. vibrate) when the robot touches the object. The degree of physical/haptic

reaction may vary depending on the pushing force exerted by the robot. Here is another example; we can use an AIBO robot dog as a controller for remote interaction with another AIBO. When the user rotates the head of the “mediator” AIBO, the remote AIBO will rotate its head correspondingly. If the mapping is designed and calibrated adequately, the gaze direction of the remote AIBO is clearly presented to the operator. However, with the traditional GUI approach, robot operators have to solely rely on the video feed and other sensor data gathered by the robot to “imagine” the current physical state of the robot. For instance, people can consistently navigating a remote robot in the wrong directions if the GUI they use fails to provide the orientation of where the robot’s camera is facing (Yanco et al., 2004).

Physical feedback provided by TUIs can range from very expressive force and haptic feedback, to a more subtle set of visual feedback provided by the TUI. For instance, TUIs themselves can be transparent objects. They can be illuminated by color lights to reflect the state of a virtual system. Although this type of feedback is subtle and not tangible, the feedback is augmented by the physical shape of TUIs.

3. *Unification of Action and Perception Space.* TUIs have the potentials of combining system input and system output together onto the same device to avoid the problem of the separation of the action and perception spaces (Sharlin et al., 2004). This is one of the major issues that the mouse, keyboard and GUI paradigm suffers from. When interacting with a robot, operators usually need to break high-level task goals into low-level discrete actions that the robot can perform, and then translate each action to a key or switch on the user interface to trigger the appropriate action on the robotic platform. The action of interacting with the interface (e.g. press a button) is separated from looking at the robot’s streaming video to confirm the correct action is performed. This separation of the action and perception space may break the flow of operators’ cognitive engagement, which in turn, may slow down the task process and increase the possibility of making mistakes. By unifying the system input and output, we can lower the overall difficulty in controlling remote robots.

4. *Interaction embodiment and physical affordance.* A carefully designed TUI can directly *afford* its usage, how people should interact with it, and the results of any actions the user may take, embodying its interaction qualities (Dourish, 2001; Norman, 1988). We are born with various skills allowing us to observe and to interact with physical objects. By taking advantage of these innate abilities, or with abilities that we are trained with from an early age, we can reduce the amount of time needed to learn how to interact with TUIs and lower confusion and error rates during interaction.
5. *Spatial Arrangement*. Unlike interaction with virtual objects, we can take advantage of our spatial ability to manipulate single or multiple TUIs and to extrapolate the spatial relationships of objects that the TUIs represent. For instance, when we play chess, by looking at the arrangement of the chess pieces, we can infer the strategy that the players are using. When interacting with multiple TUIs, the organization and the structure of these TUIs may reveal important information to users. In this situation, our spatial sense can help us to better analyse the spatial relation among these objects.
6. *Spatial Mapping*: Well designed TUIs usually provide a spatial congruent mapping between the physical objects and their usage in the digital application. (Sharlin et al., 2004) We can quantify the spatial congruence between the physical object and the digital entity using Beaudouin-Lafon’s (2000) concepts of degree of integration (DOI) and degree of compatibility (DOC). Careful design can help TUIs to achieve the appropriate DOI and DOC to better convey their usage in digital applications.

We have incorporate some of the TUI benefits explained above in our system designs (describe in Chapter 4 and Chapter 5). To better illustrate these benefits in the TUI systems we created, we will use the following table from now on as part of our system evaluation in Chapter 4 and Chapter 5:

TUI Benefits	Status
Physical I/O	Unsupported/Supported
Physical Feedback	Unsupported/Supported
I/O Unification	Unsupported/Supported
Affordance	Unsupported/Supported
Spatial Arrangement	Unsupported/Supported
Spatial Mapping	Unsupported/Supported

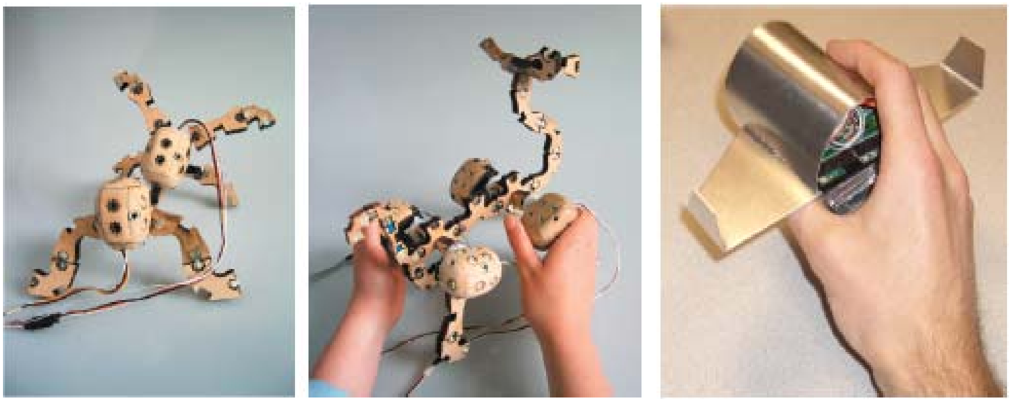
Table 3.1 Table of TUI Benefits

3.2 Early Efforts

Four early projects have demonstrated the potential of using physical objects as TUIs to manipulate robots. The Topobo toy application (Raffle et al., 2004) enables children to assemble static and motorized plastic components to dynamically created biomorphic forms (Figure 3.1 a). Further, the system replays motions created by twisting and pulling the motorized components, animating users' creations. By combining physical input and output onto the same physical object, Topobo allows kids to learn about mechanics and kinematics through rapid trial-and-error (Raffle et al., 2004)..

Another pioneering effort that utilizes a physical TUI for controlling a robot is presented by Quigley et al. (2004). They conducted a comparative study of seven mini-unmanned aerial vehicle (mini-UAV) controlling interfaces. One of their original interfaces is an airplane-shaped physical controller for directly manipulating the roll and pitch angle of a mini-UAV. (Figure 3.1 b) The result of the study showed that the physical controller achieved high ratings in control precision and responsiveness.

However, Raffle and Quigley’s research do not directly address the issues that we are looking into in this thesis. Raffle et al.’s approach of using the robot as a controller for controlling itself is interesting, but very limited. It cannot be used for remote control tasks, such as search and recue. Since the focus of his research is not on HRI, he did not provide a usability evaluation of the Topobo interface in comparison to other HRI interfaces. Quigley et al.'s approach can be considered to be a first step in our research direction. However, their project lacks validity as no real robot was used in their experiment, only a simulation. Also, the HRI tasks they used for the experiment were very limited.



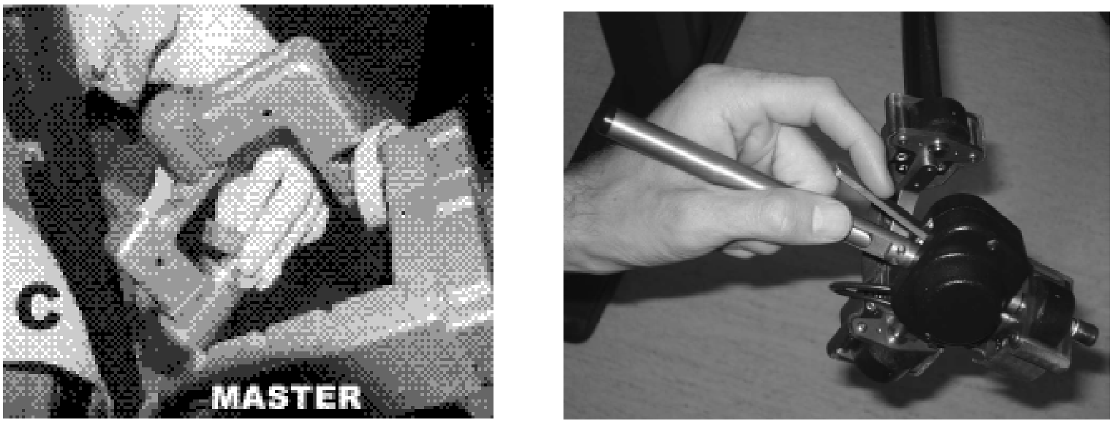
(a) A horse robot made using the Topobo system. (Raffle et al., 2004)

(b) Physical icon controller for a Mini-UAV. (Quigley et al. 2004)

Figure 3.1 – TUIs for HRI

In the field of surgical robots, two projects demonstrate the potential of using TUIs for HRI. The da Vinci telerobotic surgery system was first created in 1999 (Ballantyne and Moll, 2003). It is a telerobotic system that allows surgeons to teleoperate a set of surgery robotic arms to perform minimally invasive surgeries. To remote control the robotic arms, surgeons interact with a pair of "master" robotic arms that is capable of capturing the surgeons' 3D hand movements and translating that into robotic instrument movements (3.2 (a)).

Another surgical robot project that allows surgeons to use tangible instruments to interact with surgical robotic arms is the neuroArm project (Greer et al., 2008). Greer et al. developed a tangible user interface that allows neurosurgeons to control the hands of a surgical robot to perform neurosurgeries remotely. Their TUI is built upon a pair of PHANTOM 1.5 premium haptic devices. (Figure 3.2 (b)) This “handcontroller” interface captures the rotation and movement of surgeons’ hands to directly manipulate a surgical robot’s hands. It also provides haptic feedback directly to surgeons to recreate the sense of touch and help them navigate the robot hands intuitively. The neuroArm project well demonstrates the benefits of applying TUIs to HRI applications and matches with some of the ideas that we envisioned in this thesis.



(a) da Vinci Surgical Console (Ballantyne and Moll, 2003)

(b) neuroArm force feedback handcontroller. (Greer et al., 2008)

Figure 3.2 Surgical robot control interface.

3.3 Introducing the Ricon Concept

To follow the vision of Ishii and Ullmer (1997), Quigley et al. (2004) and the development trend in the field of surgical robots (Ballantyne and Moll, 2003; Greer et al., 2008), we intend to use a physical object as a robotic icon – *Ricon*, for controlling a robot. Depending on the application, the level of metaphor that a Ricon offers may range from “None” to “Full” as defined by Fishkin (2004)’s taxonomy. On one end of the scale (Full) proposed by Fishkin, we can use the most delicate Ricon, the robot itself for controlling another identical robot. For instance, if we are remotely controlling an AIBO robot dog, we can use another AIBO as the “controller”. To make the remote AIBO to sit up from a normal standing position, we can rotate the “mediator” AIBO’s hip joints to make it sit up. On the other end of the scale (None), we can use a physical object that does not resemble the controlled robot in anyway as a controller. In our opinion, for telerobotic interface, the metaphoric level of a Ricon should be either “noun and verb” or “full” as defined by Fishkin (2004). This is because the operator would not gain much benefit from using TUIs if the metaphoric level is lower than “noun and verb”.

For single-robot control, an appropriate Ricon should provide a tight spatial mapping (Sharlin et al., 2004) between itself and a real robot. The shape, size and weight of a Ricon should reflect the physical properties of the robot it represents. For instance, a large and heavy robot should be presented by a large and heavy Ricon. When users hold different Ricons in their hands, they can immediately feel the difference between different Ricons. In addition, it is important and beneficial if we utilize the physical constraints of Ricons to reflect the limitations of the controlled robots. One such example is that each Ricon occupies a portion of the physical space. Thus, two Ricons can never “collide into” each other. Because of this physical nature, two robots controlled by two Ricons cannot collide with each other. Thirdly, the degrees of freedom (DOF) a Ricon provides should match the DOF of a robot. For instance, if a robot is capable of moving and rotating in a two dimensional space, then we should use a Ricon that supports these functions with no more or fewer DOF. Fourthly, the level of metaphor that a Ricon affords should be determined by the task. Sometimes, there is no need to use a Ricon that is exactly identical to the controlled robot. For instance, if we want to control an AIBO to walk, it would be difficult to use another AIBO as a controller to imitate a sequence of walking posture by moving its four legs at the same time.

When interacting with a group of robots, multiple Ricons can be used. The benefits of allowing the human operator to access multiple Ricons are manifold. Firstly, users can manipulate more than one Ricon at a time. For example, a user can use both of his or her hands to manipulate many Ricons at the same time. This ability would provide an intermediate solution for the call-request problem we mentioned in the previous chapter. If one robot stopped working within a team which causes the entire team to wait, multiple

Ricons would allow the user to tackle this problem from multiple angles and solve this kind of problem more quickly.

Secondly, Ricons with different color, shape and size can be used to represent different types of robots within a group. We can use the “chess” metaphor to illustrate this problem. Each Ricon can be seen as a chess piece. The color, shape and size of each Ricon can directly reflect their usage intuitively. When these Ricons are placed on a table surface, we can think of different chess pieces are placed on a chess board. An experienced chess player can infer the strategy that each player is using and the overall balance of a chess game (who is winning or losing) by simply staring at the chess board. The same kind of benefit can be gained by looking at the placement of each Ricon on a table surface to understand the workflow among robots.

Thirdly, users can use Ricons to configure different group formations easily. Since TUIs take advantage of our innate skills of spatiality and interacting with physical objects, we can easily move, adjust and form multiple objects into any kinds of formations.

Fourthly, we can create physical tools to assist us in managing multiple Ricons. For instance, reflecting on racks used for organizing Pool or Billiard balls, we may build a triangle shaped “tie” to band multiple Ricons together to form a triangle formation. By pulling the tie, we can direct a group of Ricons to a desired location. By taking the tie off, we break the group relationship.

3.4 Possible Drawbacks

Although there are many benefits of using TUIs in HRI, there are also some drawbacks:

1. *TUIs can be difficult to generalize to handle different kinds of tasks.* Just like the physical tools that we use in our daily lives, in order to maximize their user ability in a particular area, tools are designed and created for specific tasks. For instance, rulers are used for measuring scales, and hammers for hammering nails. These cannot be used as a tool for painting pictures. There is no “universal tool” that exists in the physical world to compare to the existence of mouse in the virtual world. Thus, the nature of TUIs

prevents themselves from being as “flexible” and “portable” as a mouse. A TUI created for one type of HRI tasks may not be usable for other HRI tasks.

2. *TUIs are physical objects that must obey with physical laws*. They can be used in undesirable ways to produce unintended results. For instance, if we use a Ricon to present a robot and use it on a table to control the movement of a robot, an accidental shake to the table may knock the Ricon out of its intended placement and affects the location of the controlled robot. In the physical world, we do not have the “invalid action filtering” and “undo” mechanisms that are commonly implemented in the virtual world. Therefore, TUIs need to be designed carefully to prevent or lower the chances of accidental interactions from happening.
3. *It is hard to reflect digital information on the surface of TUIs*. Due to the limitation of technology, we do not have flexible displays that can be attached onto any kinds of surfaces. Without using projectors to project information onto TUIs, it is hard to display digital information, such as warning message, on the surface of TUIs. In telerobotics, sensors equipped on a robot can gather important information, such as temperature readings, to help the operator to better assess the remote environment that the robot is within. Without the aid of sensors, the operator would lose much “awareness” information he/she needed to complete a task. Thus, without being able to display this sensor information on the TUI’s surface can drastically lower the quality of remote robot interaction.

3.5 Conclusion

In this chapter, we explain our design motivation of applying TUIs to HRI from a theoretical point of view. We break down the benefits of applying TUIs to HRI into the following areas: *physical input and physical output, physical feedback, unification of action and perception space, interaction embodiment and physical affordance* and *spatiality*. Then, we present previous works that touched upon the idea of using TUIs for interacting with robots. Based on our experience with TUIs and literature reviews, we propose the concept of Ricons, robotic icons, and explain how they can be used for interacting with robots. Just like any other interface, TUIs have its own limitations. However, their benefits outweigh their drawbacks. Given the unique challenges posed by robotic interfaces, and the unique benefits offered by TUIs, we see a great potential in combining these research areas to develop new and novel ways to interact with robots.

Chapter 4. Exploring the Use of Tangible User Interface for Human Robot Interaction

In the previous chapter, we discussed our vision of using physical robotic icons, or *Ricons*, for controlling single and a group of robots. We suggested that by utilizing and taking advantage of TUIs, we can overcome some of the difficulties that traditional human-robot interfaces suffer from. In this chapter, we present the first study which is a comparison experiment for evaluating the usability of a generic 3D TUI for robot control tasks. In this study, we utilized a pair of Nintendo Wii™ game controllers (Wiimote and Nunchuk) as the TUIs. In order to assess the quality and effectiveness of the Wii controllers as robotic interface, we designed an experimental test bed that allowed us to test them against a generic input device – a keypad. A user study is conducted to investigate the advantages and drawbacks of each interaction method in practical HRI tasks. The design, implementation, test result and discussion are presented in this chapter.

4.1 Selecting a Generic TUI: the Nintendo Wii™ Remote

During the selection of TUIs, the Nintendo Wiimote came to our attention. The Wiimote clearly differentiates itself from other generic controllers in terms of the interaction style. Instead of pressing buttons, the Wiimote allows players to use motions such as, swing, shake and thrust to interact with the virtual objects on the TV screen. (Figure 4.1) Players feel more immersed and satisfied when using the Wiimote due to the fact that virtual entities in games react to their physical inputs. Although the Wiimote does not qualify as a highly specialized TUI, it can be categorized as a generic 3D tangible user interface due to its ability to capture physical input and to interact with digital entities. Also, due to its generality, the Wiimote offers a good spatial mapping for simple tasks that utilize arm and hand gestures. Therefore, we think the Wiimote is a good starting point for exploring the use of TUIs in HRI tasks. Success in mapping a robotic task to a tangible/gestural interaction via a simple TUI will point to the great potential of better and more elaborate TUIs in more complex HRI tasks.

Figure 4.1 – A gamer swings a Wiimote to hit a virtual baseball. (from http://cache.kotaku.com/assets/resources/2006/11/wii_baseball.jpeg)

4.2 System Design and Implementation

In order to explore the possibility of using gestures for HRI, we were looking for a robotic platform that would allow us to gain full and flexible control in lab settings. The robot should be able to response to both high level commands (such as walking or turning) and low-level commands (such as rotate a specific joint by a certain number of degrees) to match the meaning of both abstract gestures (such as arbitrary hand gestures used in a speech) and specific gestures (such as teaching others a specific movement by demonstrating a similar gesture). Moreover, we were searching for an anthropomorphic or zoomorphic robot that resembles the human skeletal structure to a degree in order to achieve an intuitive mapping between the user interface and the robot in posture control tasks. In search for robots that satisfy the above criteria, we found that the AIBO robotic dog to be a suitable platform for the studies. The AIBO is a zoomorphic robot that resembles parts of the human skeletal structure. For instance, the AIBO has “shoulder” and “elbow” joints on its forelegs which act similarly to human’s shoulder and elbow joints. By using the Tekkotsu framework (Tekkotsu, 2008), developers can gain full control over the low-level actuators, high-level body gestures and movements of the AIBO.

To evaluate the usability of gesture input for HRI in contrast with a generic input device, we have designed two interaction techniques for manipulating an AIBO in a co-located setup. One of the interaction techniques supports human gesture input through a Wiimote and Nunchuk interfaces, another input technique uses a keypad as the basis for interacting with the AIBO. In order to utilize the power of Wiimote and apply it to control an AIBO, we used a PC equipped with both Bluetooth and 802.11b wireless network adapter to act as a mediator to translate and transmit the command from the Wiimote to the AIBO. (Figure 4.2)

Figure 4.2 – Communicating between the Wiimote and AIBO through a PC

Another interface that we selected for representing the generic input device is an OQO 02 Ultra-Mobile PC (UMPC) with an onboard thumb keyboard. (Figure 4.3) The OQO 02 is a scaled down version of a regular desktop PC. It has built-in wireless network adapter that can be used to communicate with an AIBO. The OQO-based “button-press and key-to-action mapping” interaction style represents a common interaction technique in current HRI.

Figure 4.3 – OQO 02 Ultra-Mobile PC

When designing the interface we had to deal with a short (about 500ms) latency issue, resulting from the wireless network transmission and the robot’s electromechanical start up time. To maintain the fairness of the experiment, the underlying controlling code for both techniques was identical. Thus, the amount of lag the participants experienced was the same using both techniques, unless there were random wireless transmission spikes.

4.2.1 Wiimote & Nunchuk Input

The Wiimote consists of a D-pad, a speaker, four LEDs and eight buttons. It communicates with the Wii via Bluetooth wireless link. A complete 3-axis accelerometer (Analog Devices, 2007) located inside the Wiimote measures a minimum full-scale range of ± 3g with 10% sensitivity. It can measure the static acceleration of gravity in tilt-sensing applications, as well as dynamic acceleration resulting from motion, shock, or vibration. (Analog Devices, 2007) An extension port is located on the bottom of the Wiimote to allow peripherals such as a Nunchuk to be attached. The Nunchuk has an analog stick and two buttons and uses the same accelerometer on the Wiimote to support motion sensing.

In order to understand the Wiimote's motion sensing capability, we need to examine its acceleration measuring mechanism. According to the Data sheet (Analog Devices, 2007) of the ADXL 330 accelerometer:

> "*The sensor is a polysilicon surface micromachined structure built on top of a silicon wafer. Polysilicon springs suspend the structure over the surface of the wafer and provide a resistance against acceleration forces. Deflection of the structure is measured using a differential capacitor that consists of independent fixed plates and plates attached to the moving mass… Acceleration deflects the moving mass and unbalances the differential capacitor resulting in a senior output whose amplitude is proportional to acceleration. (p. 11)*"

In other words, the sensor does not measure the acceleration of the Wiimote, but rather the force exerted by the test mass on its supporting springs. (WiiLi, 2008) When the Wiimote is at rest on a flat surface the accelerometer reading is 1 g (approximately 9.8 m/s^2) due to gravity. When it is in a free fall motion, the reading is close to zero. These facts implies that it is possible to derive a relatively accurate measuring of the pitch and roll angle of the Wiimote when it is reasonably still. This is because when the Wiimote is accelerating (e.g. when a user is swinging the Wiimote), the acceleration value sensed by the Wiimote is due to the force exerted by the user rather than the pulling of gravity. Thus, the tilting angle derived based on this force does not represent the current position of the Wiimote. The Wiimote specifications also imply that the accelerometer cannot detect the rotation angle around the gravitational axis. For instance, when the Wiimote is facing up (e.g. the A button is facing upward) and resting on a flat surface, the Z-axis (Figure 4.4) of the accelerometer is parallel to the direction of gravity. Thus, it does not matter how we orient the Wiimote on the surface, the acceleration value sensed on the Z-axis always remains the same. This means that the Wiimote "loses" one degree of freedom when one of the axes of the accelerometer is aligned with the direction of the gravity.

Figure 4.4 – The coordinate system of Wiimote

Due to the constraints associated with the accelerometer and the unavailability of a complex motion analysing package, we decided to focus our efforts on measuring pitch and roll angles for recognizing arm and hand gestures. In the experiment, we wanted to allow participants to use large arm movements for controlling an AIBO. The reasoning was that large arm movements are easier to distinguish when processing the Wiimote data and relatively intuitive and easy to memorize by the user. Therefore, we decided to rely on only using the pitch angle of the Wiimote and Nunchuk to predict arm positions. In this case, we use the Wiimote and Nunchuk as a one degree of freedom input devices to measure the rotation angle of a person's elbow and shoulder joint in relation to the arm rest position.

In order to access the acceleration value sensed by the Wiimote and Nunchuk we used Brian Peek's C# library (Peek, 2007) for acquiring the accelerometer readings. To covert the raw acceleration value into rotation angles, we enter the calibrated raw acceleration values into the following equation, where the variable a_y denotes the calibrated acceleration value along the Y-axis: (WiiLi, 2008)

$$\text{Pitch} = \text{asin}\,(a_y / 1)$$

4.2.2 OQO 02 Thumb Keyboard

The generic input device that we have chosen as a non-TUI reference system is the thumb keyboard on an OQO 02 UMPC (Figure 4.3). The OQO 02 is equipped with a 1.5 GHz VIA C7M ULV CPU, 1GB DDR2 SDRAM and a 60GB HDD. It runs on Windows® XP

Tablet PC Edition 2005. It supports both 802.11 a/b/g and Bluetooth network standards. The input devices on the OQO 02 include a dedicated pointing stick, a backlit thumb keyboard with a total of 58 keys (including function keys, letter keys and a number pad) and a digital pen. The letter keys on the thumb keyboard follow the QWERTY keyboard layout. The OQO 02 can be either powered by a removable lithium-ion polymer battery or an AC charger. In the comparative study, we used the thumb keyboard solely for controlling an AIBO.

4.3 Experimental Design

To compare and better understand how well people can learn and utilize the aforementioned techniques when controlling a robot, we designed an experimental test bed based on two tasks for comparing the techniques in terms of speed, accuracy and subjective preferences of the participants. Our goal was to explore the benefits and drawbacks associated with each interaction technique, and to try to point out which technique supports a more effective, intuitive and rich user experience when interacting with a robot.

4.3.1 Pilot Study

Before the full user study, we had conducted a pilot study to debug the experimental design, test the usability of both interaction techniques and the experiment fairness under different conditions. The pilot study included 8 participants recruited from our lab. We found that our posture recognition technique does not suit well with people who have large body size. This is because when these people wear the pair of Nunchuks on their biceps, the Nunchuks usually tilt outwards, which produces “inaccurate” accelerometer readings in compare to people who have smaller body sizes. Thus, we changed the system to allow for a more flexible range of input. However, misrecognition still occurred during the pilot study. To minimize the impact of this problem on the participants’ task completion time, we modified the underlying software component that supports the interaction to automatically record the time when each posture command is triggered. The examiner also used the same software to manually log the time when a correct posture is preformed by pressing a button on a keyboard. A video tape recorder was used for backup purposes, capturing the entire session for replay and time synchronization purposes.

To enable participants to navigate the AIBO, we initially used the “W, A, S and D” key mapping on the OQO keypad for the navigation test. However, in this particular key arrangement, users only need to use their left thumb for most of the movements they need to perform. On the other hand, with the Wiimote technique, users have to use both hands with equal amount of effort to navigate the AIBO. To balance the single hand interaction technique with an asymmetric bimanual (Balakrishnan et al., 2000) interaction technique we revised the key mapping of the keypad interface (the revised mapping is explained in detail in the next section).

4.4 User Study

4.4.1 Participants

For the comparative user study, we recruited twenty participants (16 males and 4 females) from the University of Calgary; each participant was paid $10 for taking part in the experiment. Ages ranged from 18 to 29 (*M* = 21.75, *SD* = 3.05). All of the participants reported to use computer keyboard everyday. Among all of the participants, eighteen people were right-handed, one person was left-handed and one person was ambidextrous. All of the participants indicated that they have some sort of computer game experience. Fifteen participants reported to play computer games on a daily or weekly basis. Seventeen participants indicated that they “often” or “very often” use computer keyboard to play games. Six participants reported no prior experience playing the Nintendo Wii. Out of the fourteen people who had previous experience with the Wii only three participants reported to play it on a weekly basis. The other 11 indicated playing either “Monthly” or “Rarely”.

4.4.2 Task and Procedure

The experiment was designed for two different tasks, *robotic navigation* and *robotic posture*, each with two difficulty levels. The participants were asked to perform both tasks with both interaction techniques. Thus, in total, participants had to go through four sub experiments in order to complete the study. The order of techniques was counterbalanced among participants by alternating the tasks order, thus ten participants started with the Wiimote Interface and ten participants started with the OQO interface. The experiment was conducted following a written protocol (see Appendix A.3 for details). Participants were asked to start with one interaction technique to complete both *navigation* and *posture* tasks and then switch to the other technique and repeat the two tasks. During the experiment, each participant was asked to complete four sets of questionnaires after each trial and, once finished, to complete a post-study questionnaire which was followed up with a non-structured interview. Each study took around 60 minutes to complete.

To allow participants to learn and practice each interaction technique and to familiarize themselves with the tasks a practice trial was administrated before the full experiment started. The administrator demonstrated the interaction techniques and presented guidelines on how to complete the tasks. Then, the participants would try out the interaction technique until they felt proficient and comfortable to move on.

The main dependent measure in the experiment was the task completion time. In addition, we recorded the number of errors that the participants made with each interaction technique.

4.4.3 Task 1 – Navigation

In this task, the participants were asked to navigate the AIBO through an obstacle course (Figure 4.5). The obstacle course is 262 cm in length and 15.3 cm in width. The goal of this test is to see how well both interaction techniques support user control in a fairly delicate robotic navigation task. Eight different navigation actions were provided to the users: *walk forward, stop, walk forward while turning left, walk forward while turning right, rotate left, rotate right, strafe left and strafe right.*

To motivate the participants to use all actions, we designed two routes based on the same obstacle course shown in Figure 4.5 for the task. For the easier route, participants were not forced to use any particular action during the course of the obstacle course and were allowed to use any combination of actions they want. However, for the harder route, participants were forced to use rotation and strafing in addition to walking and turning in order to complete the obstacle course successfully. A dotted yellow line on the course (Figure 4.5) indicated the starting point of the strafing action. The solid yellow line indicated the starting point of the *rotate right* action. In order to finish this task, the participants were asked to complete the easier route first followed by the harder trail.

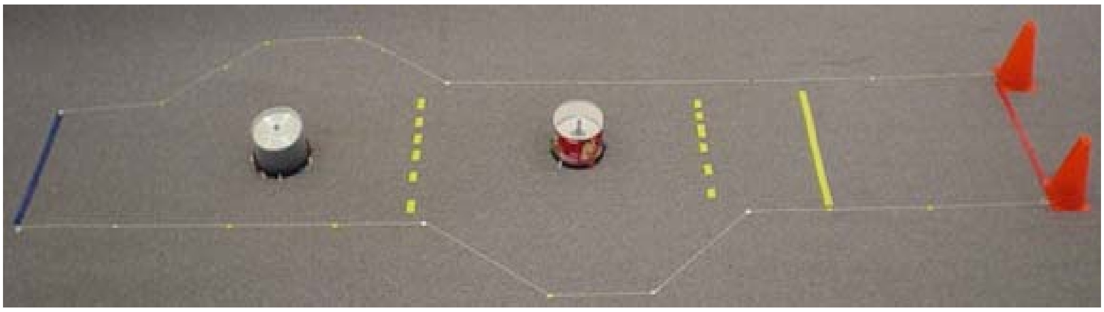


Figure 4.5 – The obstacle course

Before the start of the experiment, the test administrator would remind the participants to complete the obstacle course as fast as possible, and try to make as few errors as possible. An error in this task is defined as hitting obstacles, navigating the AIBO out of the route boundary or failure to perform required actions at the specified locations. If a participant navigated the AIBO out of the route boundary, then she/he had to navigate it back to the route and continue on. If a participant failed to perform the required action at certain locations during the trial the examiner had to physically move the AIBO back to that location and ask the participant to try again. This error correction mechanism could have introduced a variable amount of time into the task completion time depending on how fast the examiner moves the AIBO back to the right location. We emphasized the implications of this penalty to the participants. We were pleasantly surprised to see that none of the experimental trials required the test administrator to physically move the AIBO or to manually correct any out-of-bound navigation errors.

The Wiimote interaction mapping used for the navigation task is presented in Figure 4.6 and the mapping for the keypad interaction is presented in Figure 4.7. The gesture mapping for the Wiimote controller was developed based on horseback riding techniques metaphor. The participants were told to think of the pair of Wiimotes as a neck rein on the AIBO. For instance, pulling both Wiimotes back will stop the AIBO; pulling the right Wiimote only will rotate the AIBO to the right, etc.

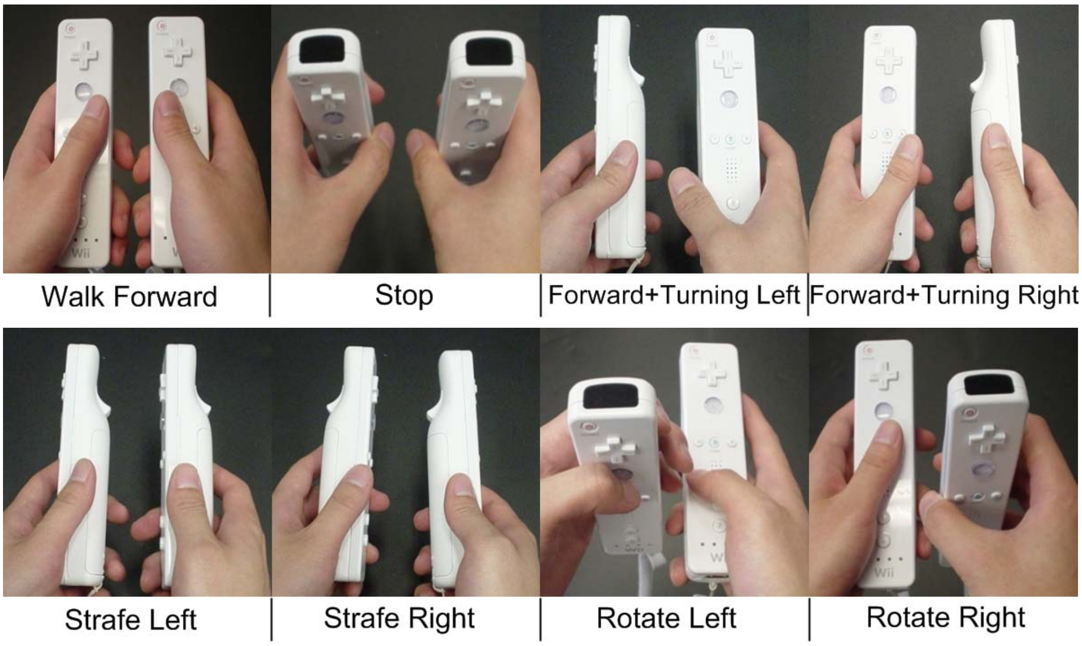


Figure 4.6 - The Wiimote interaction technique for controlling the movement of the AIBO

Figure 4.7 – Key-Movement Mapping

Due to the nature of the task, the gesture-to-robot action mapping is somewhat indirect. In this case, the participants are not controlling a single joint of the AIBO but rather the spatial kinematics of the robot when navigating it through the obstacle course. This implies a non-ideal "degree of integration", and a weaker "degree of compatibility" (Beaudouin-Lafon, 2000) for the gesture-to-robot action mapping in this task. However, we can argue that the horseback riding metaphor provides efficient and intuitive mechanism for dealing with this abstract mapping.

The keypad-to-robot action mappings are: Forward – W + 2, Stop – S + 5, Forward + Turning Left – A + 2, Forward + Turning right – W + 6, Strafe Left – A + 4, Strafe Right – D + 6, Rotate Left – S + 2, and Rotate Right – W + 5. (The plus sign means pressing and holding the keys on both sides of the sign).

The Data collected from this task was analyzed using a 2 x 2 within-subjects ANOVA for the following factors:

Technique: Wiimote, Keypad

Difficulty: easy, hard.

4.4.4 Task 2 - Posture

This task is used to examine the usability of both interaction techniques for low-level robot control. In this task, we asked the participants to control twelve different postures with the forelegs of the AIBO. The test administrator would show an image of the AIBO with a posture on a computer screen. Then the participants were asked to control the AIBO (Figure 4.10) so it will imitate the posture (Figure 4.9) presented on the screen.

In the experiment setup, we have pre-defined four different postures for each foreleg of the AIBO. (Figure 4.8) We selected ten postures out of the sixteen possible combined postures using both forelegs. Then, we divided them into two groups of postures which can be chained together to create gesture sequences (Figure 4.9). The only difference between these groups of postures is that in order to transform from one posture to another within a group, the participants have to manipulate either one foreleg or both forelegs of the AIBO to complete the transition. We define the group of postures that require only one arm movement during the transition as the easier set, and the other group as the harder set. For the experiment, the participants were asked to perform the easier set first followed by the harder set.

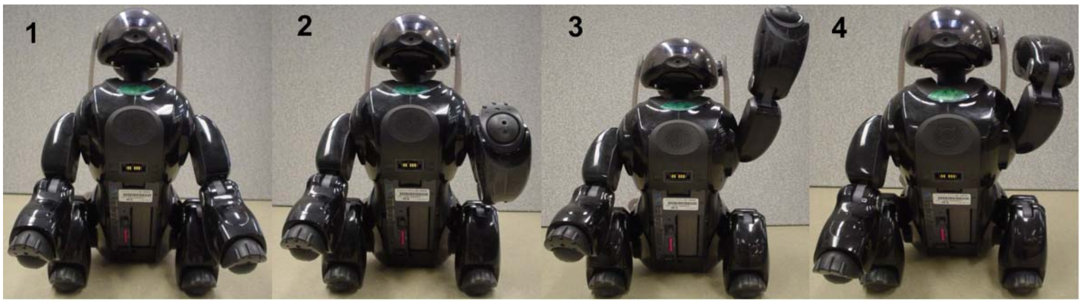


Figure 4.8 – The possible postures for each foreleg of the AIBO

Figure 4.9 – Posture 1-6 is the easier posture group. Posture 7-12 is the harder posture group.

Similar to task 1, we measure the task completion time and the number of errors. The task completion time in this task is defined as the time that elapsed since a new posture image was displayed on the screen till the time the participants completed the correct robotic posture. Completion time was measured automatically by the software according to the user sensed gestured, with a manual measurement for backup. The error in this case is defined as performing a posture that is different from the posture displayed on the screen. If a participant fails to perform the correct posture, then he/she needs to correct themselves. The time it takes the participants to think and correct their postures is also taken into account as part of the task completion time. Since the harder posture set requires the participants to move both forelegs of the AIBO, the actions can be preformed either sequentially or simultaneously. In this case, we did not constrain the participants to any of the input styles, allowing them to gesture either sequentially or simultaneously, as long as they feel it is the fastest and most intuitive way to complete the postures.

For this task, the function mapping for the Wiimote interface is presented in Figure 4.10 and the mapping for the keypad interface is presented in Figure 4.11.

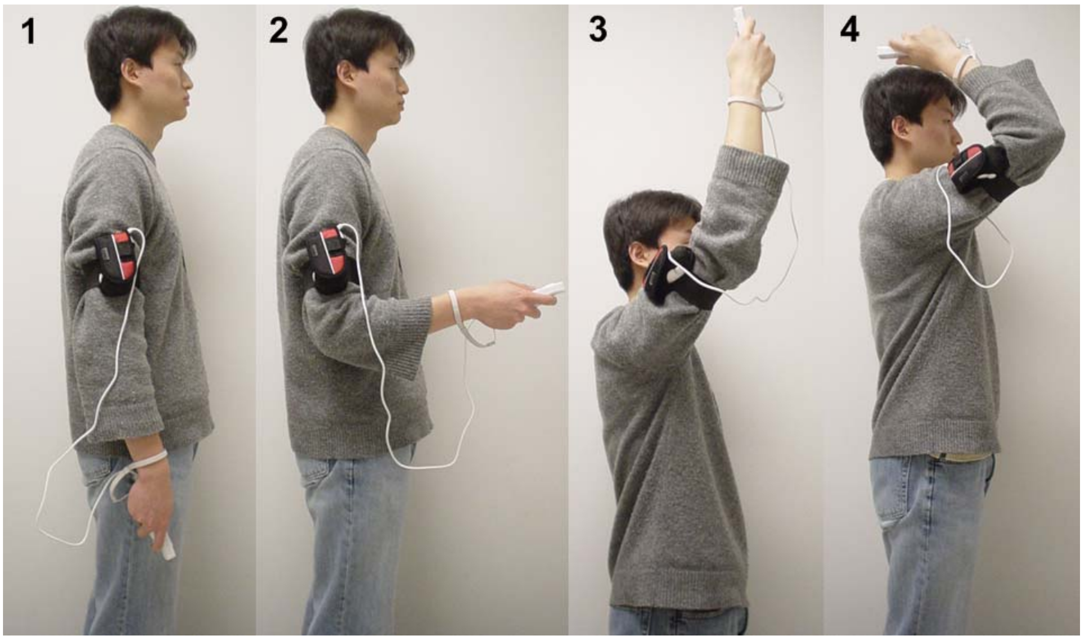


Figure 4.10 – Arm postures input. Theses postures correspond to the four AIBO postures illustrated in Figure 4.8

Figure 4.11 – Key-Posture Mapping.

For the gesture input technique, the participants directly adjust the position of the forelegs of the AIBO using their own arms. Compare to the navigation task, the gesture-to-robot action mapping in this case has an almost prefect degree of integration, and a high degree of compatibility (Beaudouin-Lafon, 2000).

For the keypad-to-robot action mapping, the four letter keys on the OQO control the right foreleg of the AIBO. The four number keys control the left foreleg of the AIBO. By pressing either X or 8, the AIBO will perform Posture 1 (Figure 4.8) with either its right foreleg or left foreleg. By pressing either Z or 9, the AIBO will perform Posture 2. By pressing either A or 6, the AIBO will perform Posture 3. By pressing either Q or 3, the AIBO will perform Posture 4 (Figure 4.8).

The Data collected from this task was analyzed using a 2 x 2 within-subjects ANOVA for the following two factors:

Technique: Wiimote/Nunchuk, Keypad

Posture: posture 1 to 12 (Figure 4.9).

4.5 Results

4.5.1 Task 1 – Navigation

4.5.1.1 Task Completion Time

A 2 x 2 (*Technique* X *Difficulty*) ANOVA, with repeated measures on both factors, revealed no significant *Technique* X *Difficulty* interaction ($F_{1,19} = 1.54$, $p = 0.23$), which suggests that performance with the techniques is not substantially influenced by the difficulty level. There was a significant main effect for *Technique*, $F_{1,19} = 12.19$, $p <.001$, indicating that overall task completion time for the Wiimote technique (M = 43.2s, SD = 6.9s) was 10% faster than for the keypad technique (M = 48.5 s SD = 6.7s) (Figure 4.12). As we expected, the main effect of *Difficulty* was significant, $F_{1,19} = 115.61$, $p < .001$, with the mean jumping from M = 38.7s, SD = 4.6s for the easy trail to M = 53.0s, SD = 8.1s for the hard trail.

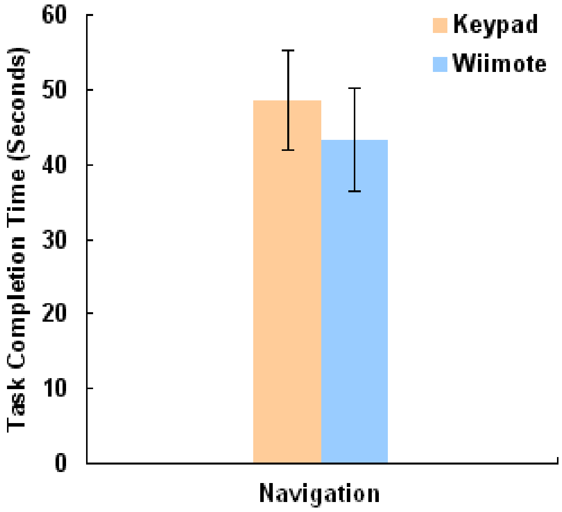


Figure 4.12 – Mean Task Completion Time for Navigation Task.

4.5.1.2 Number of Navigational Error

A two-way ANOVA was used to determine if there were differences on the number of errors (dependent variable) participants made using the Wiimote and keypad techniques when performed the navigation task under different difficulty levels. The result of the ANOVA showed no significant *Technique* X *Difficulty* interaction ($F_{1,19} = 0.03$, $p = .87$), which suggests that the number of errors made using different techniques is not significantly influenced by the difficulty level. There was a significant main effect for *Technique*, $F_{1,19} = 9.81$, $p < .01$, indicating the errors that participants made using the Wiimote technique ($M = 0.35$, $SD = 0.4$) is 43% less than using the keypad technique ($M = 0.83$, $SD = 0.6$). The result also showed a marginally significant main effect for *Difficulty* ($F_{1,19} = 3.96$, $p = .06$), with mean varying from $M = 0.43$, $SD = 0.4$ for the easy trail to $M = 0.75$, $SD = 0.6$ for the hard trail.

4.5.2 Task 2 – Posture

4.5.2.1 Task Completion Time

A 2 x 12 (*Technique* X *Posture*) ANOVA on the task completion time for the posture task showed a significant *Technique* X *Posture* interaction effect ($F_{11,209}$ = 8.43, $p < .001$), which means that the *Technique* effect varies with *Posture* or vice versa.

On the average, there was a significant effect for *Technique* ($F_{1,19}$ = 67.37, $p < .001$), with mean times reducing from 2.2s (*SD* = 0.4s) with keypad, to 1.5s (*SD* = 0.3s) with Wiimote/Nunchuk; On the average, a 32% reduction in task completion time between the two conditions. On the average, pairwise comparisons showed that there was a significant difference ($p < .05$) between the techniques for posture 1, 2, 7, 8, 9, and 10. But, there was on significant difference for the other postures. (Figure 4.13) Also, on the average, the test showed a significant effect for *Posture* ($F_{11,209}$ = 27.77, $p < .001$).

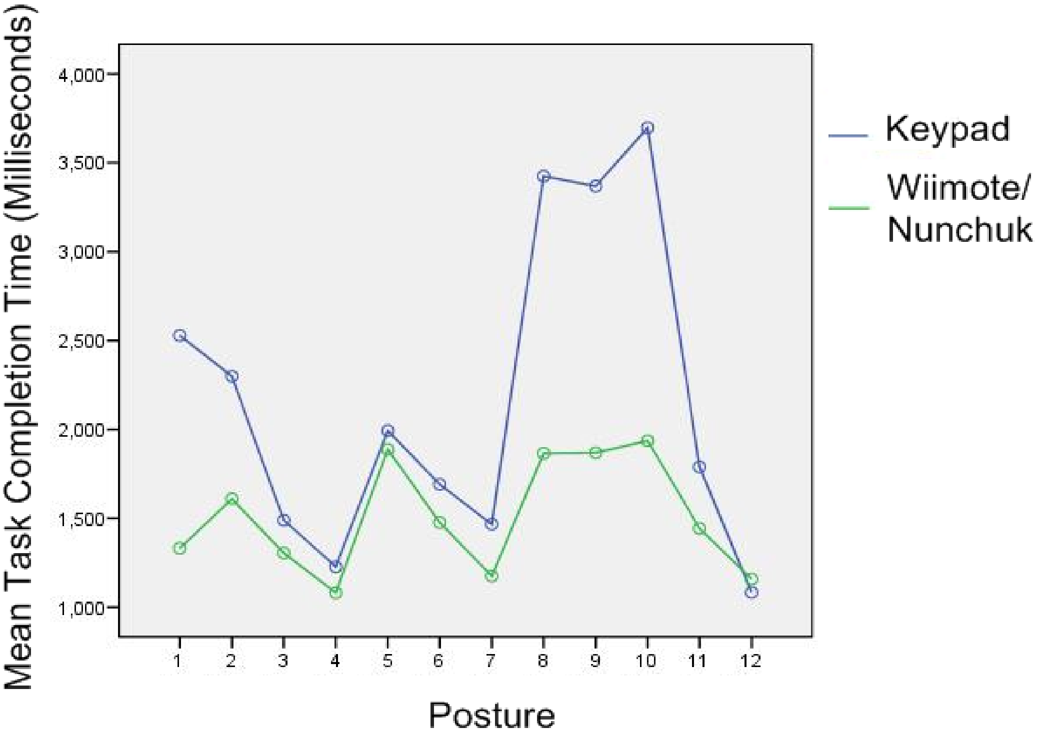


Figure 4.13 – Pairwise comparisons of the mean task completion time for each interaction technique according to posture.

4.5.2.2 Number of Posture Error

For the keypad interface, participants had made 1.5 (*SD* = 1.2) errors on average for both difficulty levels. However, none of the participants had made any errors using the Wiimote/Nunchuk interface. As anticipated, a paired t-test showed a significant difference ($t_{19} = 7.44$, $p < .001$) between the techniques.

4.5.3 Discussion

The results presented in the previous section point to the Wiimote and the Wiimote/Nunchuk interfaces outperforming the keypad interface in terms of task completion time in both the *robotic navigation* and the *robotic posture* tasks. The differences between the interfaces, although statically significant, are a little underwhelming in their magnitude.

When attempting to explain this for the navigation task, we should consider that both interaction techniques use a set of abstract key and gesture combinations to represent specific robot movements. Since none of the participants have prior experience with these input methods, they have to learn and memorize the mappings of both techniques in order to navigate the AIBO. This abstract mapping between the user interface and the robot action added an extra layer of cognitive load for the participants to process during the experiment. Although pressing buttons should not be slower than performing gestures, the study showed that the participants finished the obstacle course quicker with TUIs input than with button input. We believe that although both interfaces require the participants to think about the abstract mapping before carrying out any actions, the TUI provides a slight advantage.

When using the Wiimote, participants do not need to focus on their hands while performing a posture. They are naturally aware of the spatial location of their hands. For the keypad interface, we observed that the participants have to constantly shift their attention back and forth between the keypad and the AIBO to look for the buttons they want to press and to confirm if they triggered the intended action. The consequences of shifting attention constantly between the interface and the AIBO may result in action overshoot (for example, overturning a corner) and can break the continuity of the task when participants have to stop the AIBO before they decide which action to take for the next step. This practical separation of action and perception spaces in the non-TUI approach (Sharlin et al., 2004) is perhaps the reason for the slower task completion time when using the keypad.

Another possible reason for the faster task completion time when using the Wiimote/Nunchuk in the navigation task may be the zoomorphic rein-like mapping we used. While the mapping offered in this condition is not ideal (see our previous discussion of its degrees of integration and compatibility) the mapping does afford a simple, and arguably intuitive interaction metaphor.

Although the study results indicate that gesture input is faster for the navigation task, we are not suggesting it would always be a better solution than button input for this type of tasks. As we mentioned earlier in the pilot study section, the keypad mapping that we used was arguably not the most intuitive mapping we can come up with. A “W, A, S, D” key configuration would probably be more intuitive to use since it requires less key combinations and is a commonly used mapping in computer games for navigational tasks. However, we believe that the test results demonstrate that when participants are limited to use asymmetric two-hand interaction techniques to control a robot, TUI-based input tends to be more intuitive to use than button input.

For the navigation tasks, we did not expect that there would be a significant difference between the numbers of errors participants made using the different techniques. However, the data showed the opposite. Participants made 43% more errors with the keypad interface than with the Wiimote interface. Many participants felt that this was due to the small key size and the less intuitive mapping between buttons and robot actions.

For the posture tasks, we can see that on average there was a significant difference in task completion time between the postures that required two arms movement and the ones that only required one arm movement. By observation, we found that when the participants were using the TUI, they were extremely engaged and focused on the computer screen that displayed the posture images. However, when the participants used the keypad interface, they often looked at the computer screen first, and then focus on the keypad to find the right button to press. This attention shifting problem slowed down the participants’ task completion time and can again be associated with the separation between action and perception space created by the keypad.

Most participants felt they were simply mimicking the postures on the computer screen when using the TUI, but they felt the keypad interface required them to “act”. Following, we believe that the intuitiveness of gesture input had definitely reduced the cognitive load of associating user inputs with zoomorphic robotic actions.

In addition, TUI-based input tends to support simultaneous input compared to button input. As one of the participants commented, “we could do both hands (both arm movements) at the same time without a lot of logical thinking (with the Wiimote/Nunchuk interface), where with the keyboard we had to press one (button) and the other (button) if we was doing two hand movements at the same time. Although they would be in succession, they would not be at the same time.”

It is worth to point out that even though posture 1 and 2 only required single arm movements, there was a significant difference between the task completion times of both techniques. In our opinion, we think this is perhaps due to the participants not being fully trained at the beginning of the study. Thus, they tend to make more mistakes with the first few postures. This may also imply that the Wiimote/Nunchuk interface was easier to learn compared to the keypad interface and can be utilized faster.

4.5.3.1 Subjective Rating

After the study, a set of post-test questionnaires was given to the participants for them to rate the intuitiveness of both input techniques and indicate their preference. Figure 4.14 and 4.15 shows the results of participants’ ratings.

Q1	I found the keypad technique is easy to learn for the Navigation task.
	I found the Wiimote technique is easy to learn for the Navigation task.
Q2	I had difficulty remembering how to perform certain movements with the keypad technique for the Navigation task.
	I had difficulty remembering how to perform certain movements with the Wiimote technique for the Navigation task.
Q3	I found the keypad technique is easy to learn for the Posture task.
	I found the Wiimote technique is easy to learn for the Posture task.
Q4	I had difficulty remembering how to perform certain postures with the keypad technique for the Posture task.
	I had difficulty remembering how to perform certain postures with the Wiimote technique for the Posture task.

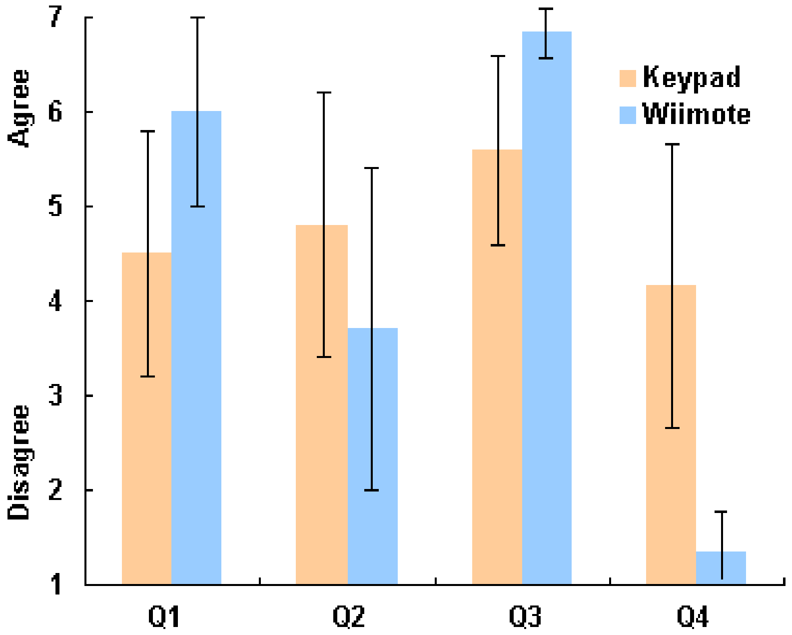


Figure 4.14 – Mean ratings on post-study questionnaire. The rating scale ranges from 1 (strongly disagree) to 7 (strongly agree).

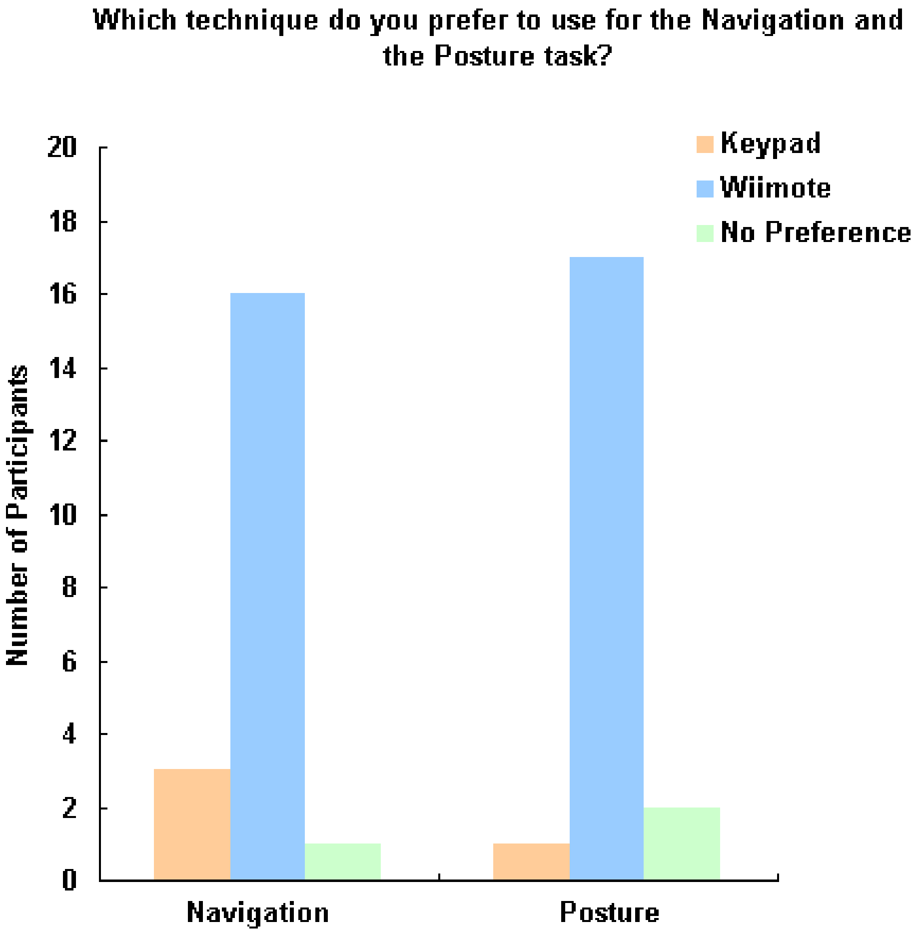


Figure 4.15 – Participants’ preference for each interaction technique.

We asked the participants who preferred to use the keypad for the navigation task about their subjective reasoning. All of them responded that they are more familiar with the keypad interface because of related computer game experiences. However, their performance indicates they completed the navigation task when using the keypad slower than when using the TUI interface. One of the participants commented, “we have to think harder when we use the keyboard, and this kind of mental overhead coupled with the lag time just makes it feel harder.”

For the participants who preferred to use the keypad for the posture task, their reasoning was that they can easily memorize the key-action mapping since there were only four postures for each arm and the buttons associated with both arms are symmetrical on the keypad layout. As one of the participants stated, “With so few postures available, the keyboard was just as easy as the Wiimote.” We agree with this participant’s comment. We believe that if the participants were well-trained with the keypad interface, they would eventually outperform the TUI in terms of task completion time. However, we think that the gestural TUI control method would prevail if the number of degrees-of-freedom and postures increase to an amount that participants cannot easily memorize, or if we deal with an interaction task that cannot afford intensive training.

During the experiment, many participants asked whether the Wiimote interface supports gradual motion sensing. The consensus indicates that people expect gesture interface to be capable of sensing and reacting to gradual changes of body motion. However, in order to maintain the fairness of the comparative study, we implemented the Wiimote interface as a state machine to match the limitation of the keypad interface.

4.5.3.2 TUI Evaluation

We also evaluated our system based on Table 3.1 presented in Chapter 3:

TUI Benefits	Status
Physical I/O	Unsupported – Physical actions performed on the Wiimotes and Nunchuks do not cause the AIBO to produce corresponding physical reactions.
Physical Feedback	Unsupported – The Wiimotes and Nunchuks do not provide physical feedback to their holders.
I/O Unification	Unsupported – The users’ action space is separated from their perception space. While manipulating the Wiimotes, users have to focus on the robot to confirm the correct action is triggered. Even though this problem applies to both the Wiimote and keypad interface, the Wiimote interface provides a much intuitive and easier mapping which does not require the users to focus on the input

	device while controlling the robot.
Affordance	Supported – We believe the Wiimotes support the horse rein metaphor well in the navigation task. However, for the posture task, there is minimal affordance that Wiimote offers. But as explained earlier, we believe users would perceive a different kind of experience if they complete the posture task with pure hand and arm gesture capturing system.
Spatial Arrangement	Unsupported – The spatial arrangement of the Wiimotes and Nunchuks do not provide awareness information to the users.
Spatial Mapping	Supported – For the navigation task, we used an abstract mapping between the Wiimote states and the AIBO’s actions. In this case, the spatial mapping is not obvious but still exists. An experienced user can tell the action of the robot is performing by looking at the state of the Wiimotes. For the posture task, we achieved a very good spatial mapping between the Wiimotes and Nunchuks states with the AIBO’s gestures.

Table 4.1 - TUI Evaluation.

Although the Wiimote interface dose not afford all of the benefits that TUIs could provide, we still observed a significant difference between the Wiimote interface and the keypad interface in terms of task completion time and the numbers of errors. This comparative study points out the great potential of applying TUIs to HRI.

4.6 Conclusion

In this chapter, we showed our design and implementation of a new interaction technique which utilizes simple generic 3D TUIs (based on the Nintendo Wiimote and Nunchuk) to capture human arm and hand gesture input for human-robot interaction. To evaluate this technique, we designed and implemented another keypad based technique and conducted a comparative user study between the two interfaces. Two tasks were used for the study: the posture task utilized a direct mapping between the TUIs and the robot, and the navigation task utilized a less direct, more abstract mapping. The result of the experiment provides some evidence that a gesture input scheme with tangible user interfaces can outperform a button-pressing input design for certain HRI tasks. We have observed a significant decrease in both task completion time and the number of mistakes participants made for both the navigation and posture tasks. The follow-up questionnaire revealed that a significant majority of the participants chose the TUI as their preferred technique for controlling an AIBO in both tasks.

Chapter 5. Touch and Toys: Interaction with a group of robots

In the pervious chapter, I demonstrated the first system that I have created which utilizes tangible user interfaces for controlling a single robot. One of the significant results from the study was that gesture-based tangible user interfaces can reduce the task completion time and the numbers of errors the users make for certain HRI tasks compared to a button-based interface. This result provides us a strong indication that TUIs have potential to improve efficiency and ease of use if mapped correctly to various HRI tasks. Therefore, I decide to take one step further to explore the possibility of using TUIs for controlling multiple robots.

In this chapter, we present two innovative interfaces that allow a single operator to interact with a group of remote robots. Using a tabletop computer the user can configure and manipulate groups of robots directly by either using their fingers (touch) or by manipulating a set of physical toys (tangible user interfaces). The primary focus of this work is on the TUI interface we created. The touch interface serves as a reference system for evaluating the usability of the TUI interface. We recruited participants to partake in an extensive user study that required them to interact with a small group of remote robots in simple tasks. The findings are presented as a set of design considerations in this chapter.

5.1 System Design and Implementation

Interaction with a remote team of robots in real time is a difficult HRI problem exacerbated by the complications of unpredictable real world environments, with solutions often resorting to a larger-than-desirable ratio of operators to robots. We implemented a TUI based interface and a touch based interface that allow a single operator to remotely control a group of robots. Both interfaces utilize a tabletop computer as the primary display for providing feedback about each robot’s status. we choose to use a tabletop computer instead of a normal monitor is because the large, public workspace surface provides a unique interaction environment that emphasizes collaboration, planning, organizing, and other spatially-situated activities (Mandryk et al., 2002; Rogers and Lindley, 2004; Scott et al., 2003), which is well-suited to the task of controlling a team of robots.

The tabletop PC that we used is a standard PC with four video outputs that combine to form a high-resolution (2800 x 2100 pixel) display surface projected onto a 146 cm x 110 cm SMART board, which also offers touch input. The TUI interface is accomplished by using a Vicon object-tracking camera system to track the location and orientation of the TUIs on the tabletop surface (see Figure 5.1(a)). Similarly, a second Vicon system is used to track the robots and report their locations back to the controlling PC, which commands the robots via 802.11 wireless and Bluetooth (see Figure 5.1(b)). We use two Sony AIBO Robotic dogs (one white one black) and an iRobot Roomba as our robots. (For more implementation detail, please refer to Appendix B.4)

The basic design of our interfaces enables the user to specify a target location and orientation for a given robot, with the system showing the user the actual current robot location. The target location is represented by an interactive icon in the touch case, or a physical toy object in the TUI case, and a line is drawn from the current location to the target to specify the robot's planned movement trajectory. The robots move in a straight line to their target location as defined by either the toy or touch interfaces. When the physical robot has reached the target location, the target icon or TUI is highlighted by a green halo (Figure 5.1(c)).

(a) The tabletop workspace with TUIs on top and the Vicon ceiling setup.

(b) The robot workspace with Vicon cameras and robots.

(c) The TUI interface. The green halo around the black dog means the black AIBO has reached its target. The white AIBO icon represents the physical robot’s location, attempting to follow the line toward the target location defined by the white-dog toy.

Figure 5.1 - Interface Overview

5.1.1 Tangible User Interface

Our goal is to enable users to intuitively associate a given TUI to a particular robot and to naturally know how to move and use the TUI without training. We used plush dogs, black and white, to respectively represent the AIBOs, and a white Frisbee to represent the white Roomba (Figure 5.2). Moving and rotating these TUIs is as intuitive to a user as any physical object, and the spatial mapping between the TUI states and the robots is direct. As an aside, the plush design of the dog TUIs makes them a pleasure to touch and comfortable to use, an important aesthetic point that we believe could add to the experience of using the TUI.

Figure 5.2 - Our tangible user interfaces and corresponding robots

We carefully selected the size of the TUIs to be similar to the actual robots and the dimensions of the physical robot space to match the tabletop. This enables users to rely on the intuition provided by the TUI dimensions, for example, two robots cannot be placed at the same location because the TUIs collide. This provides a physical constraint to the interface that reflects the real constraints of the remote robots.

5.1.2 Touch Interface

We implemented a touch interface and used it as a reference system for comparing and contrasting with the TUI interface in terms of usability. We selected a very simplistic approach where each robot is represented by a single icon. To move the icon, the user could either translate it by touching the center circle of the icon and moving it, or by selecting outside the circle and using RNT (Rotate'N Translate) a technique that enables the user to

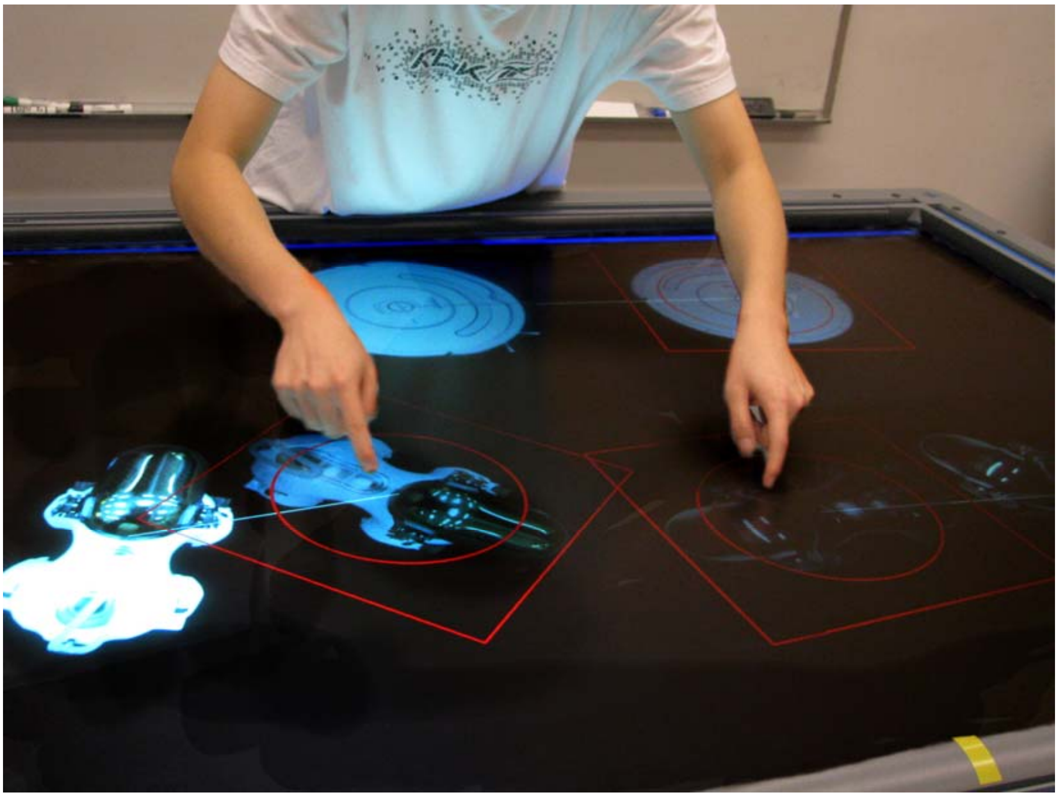


Figure 5.3 - A user simultaneously interacting with two robots. Touching inside the circle does a translation, touching outside the circle (but inside the square) performs an RNT operation.

rotate and translate the object simultaneously using only a single touch point of input (Kruger et al., 2005; Hinrichs et al., 2006) (Figure 5.3).

5.2 Evaluation

A core problem with evaluating human-robot interfaces generally, and interfaces for a group of robots specifically, is validity. People who interact with groups of robots in practice will conceivably be trained professionals dealing with real, meaningful tasks. Unfortunately, real world robotic-group users who are engaged with real tasks are very rare and often inaccessible, and simulating valid in-lab scenarios with limited off-the-shelf robotic technology can be very difficult.

We explicitly avoid this problem by focusing on the interface itself rather than the application of the interface to a task. We want to evaluate directly how people approach, respond to, and use the interfaces that we have created. While the dynamics of interaction will change with the task and training of professional operators, we feel that many of the basic interface principles and gains, the visceral level of interaction (Norman, 2004) and many usability principles of the interface itself, will remain the same. We approach the evaluation of our system with primarily qualitative techniques.

5.2.1 Experimental Design

We recruited 23 participants, aged 19-47 yrs (avg 25.5 yrs, SD 6.5 yrs), 15 male and 8 female, from the university population to participate in the study. Each participant was paid $10 per hour for their time (most took 1.5 hours and were paid $15). 20 were right handed, 1 left handed and 2 ambidextrous.

Throughout the experiment, we presented the user with a robot configuration using cut-out robot pictures on a white board (See Appendix B.2 for more detail). Following, the user was asked to put the robots into the configuration and locations that we presented to them (Figure 5.4). This was done in three stages, a one-robot, two-robot, and three-robot stage.

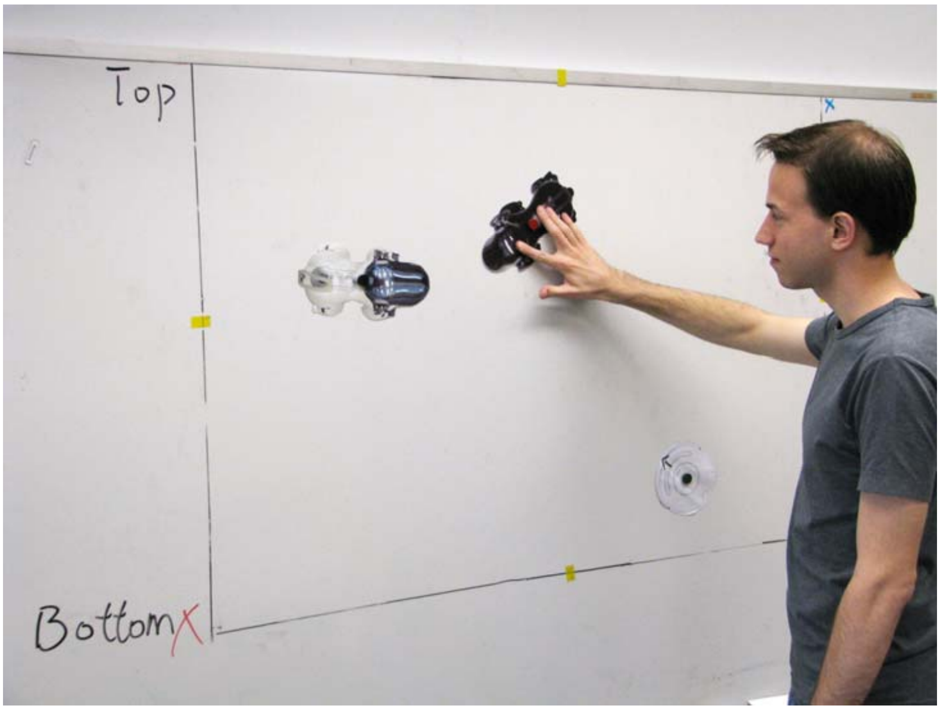


Figure 5.4 - A study administrator presenting a target robot configuration to a participant.

For each stage, the robots were set to a starting position and users were asked to move the robots to five configurations using both the touch and the toy interfaces in turn. The configurations were the same across interfaces, but changed with the number of robots. For the one-robot case, the user did the task for each the AIBO and the Roomba, for the two-robot case we used a single AIBO (white) and a Roomba, and for the three-robot case we used two AIBOs (one black, one white) and a Roomba. The order that we presented the touch and toy interfaces, as well as the order that the robots were presented in the one-robot case were counterbalanced across users, but all users were presented with the one, two, and three-robot cases in order. The user was asked to complete questionnaires before the study, after each stage and interface type, post-study, and then to go through a final interview.

5.3 Results and Analysis

Users unanimously reported (100%) that they found the graphical feedback on the table easy to understand and that it was not unnecessary, and we found no correlation between the sex, age, handedness, or past experience of the participant and their reaction to the system.

5.3.1 Task Completion time

In the one-robot case, a 2 x 2 (*Technique* X *Robot*) ANOVA was conducted. The analysis revealed no significant *Technique* (toy or touch) x *Robot* (AIBO or Roomba) interaction ($F_{1,22} = 0.15$, $p = 0.7$), which suggests that performance with the techniques is not substantially influenced by the robot type. There was no significant main effect for *Technique* ($F_{1,22} = 0.54$, $p = 0.47$). However, there was a significant main effect for *Robot* ($F_{1,22} = 19.15$, $p < .01$), indicating that the task completion time for the Roomba (*M* = 131.8 s, *SD* = 10.34 s) was 11% faster than the AIBO (*M* = 147.28 s, *SD* = 21.43 s).

In the two-robot case, a paired-t test was conducted and it showed a significant difference between the touch and toy method ($t_{22} = 2.61$, $p = .02$). With the toy interface, the participants completed the task (*M* = 170.26 s, *SD* = 26.19 s) 10% faster than with the touch interface (*M* = 188.22 s, *SD* = 32.33 s).

In the three-robot case, a paired-t test showed no significant difference between the two interaction methods ($t_{22} = 1.24$, $p = .23$).

5.3.2 Usability

We asked four ease-of-use questions (via questionnaire) after each interface type and across all three robot cases (6 times in total). The results are shown in Figure 5.5 which shows the percentage of positive responses (> 4 on a 7 pt Likert) to each question respectively. On a finer granularity, when toy and touch received a similar amount of positive response, toy received significantly more *strongly positive* responses than touch. For example, responses to the "precise control over robot movement" question in Figure looks similar across cases, but the strongly positive responses for toy/touch were 30%/7%, 30%/9%, 22%/9% for the one, two, and three-robot cases respectively.

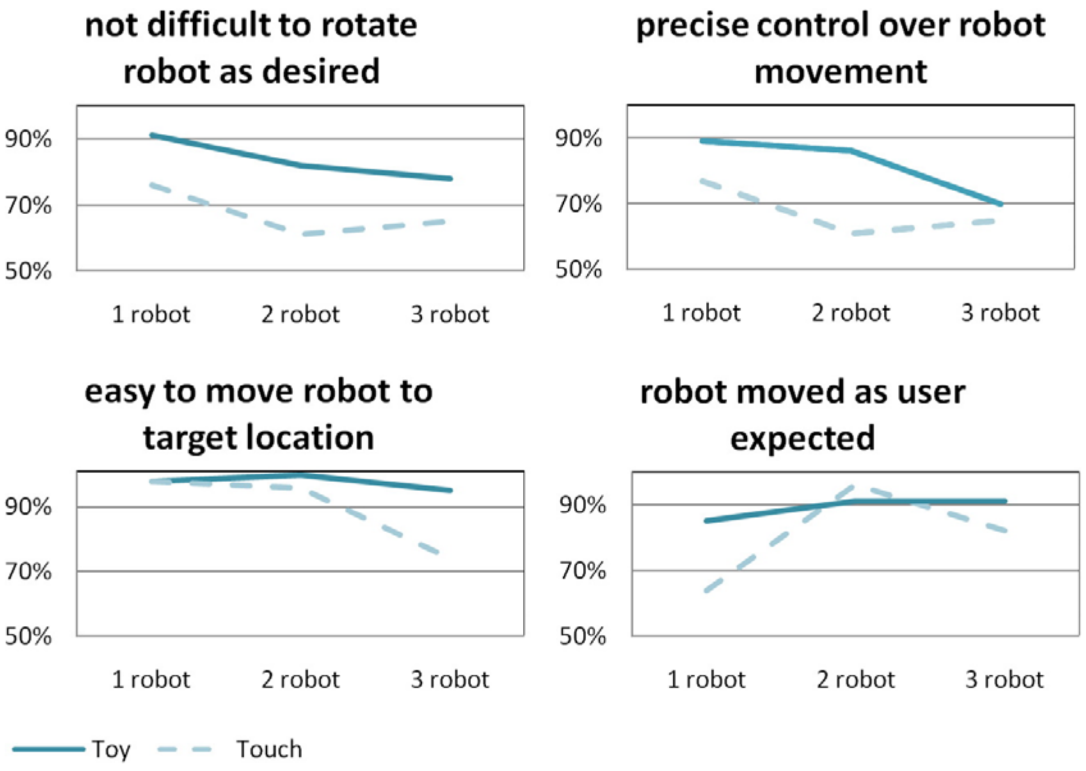


Figure 5.5 - Ease-of-use responses

Users reported that (in comparison to touch) the toy interface gives more precise control over robot movement, and makes it easier to move the robot to the target location and rotate the robot as required. Further, in the two-robot case users said it was not confusing to monitor the two robots at the same time (70% toy, 61% touch) but easy to control the robots simultaneously (78% toy, 57% touch). With the three robot case, users also said it was generally not confusing to monitor all three robots at once (70% toy, 52% touch) and that they found it easy to form the group formations asked (83% toy, 57% touch). Further, Table 5.2 reports the percentage of users that responded positively to questions about using both hands and controlling multiple robots simultaneously using the touch and toy interfaces. The table shows that users found it much easier to control two and three robots simultaneously with the toy interface than the touch interface.

5.3.3 Preference

For each of the one, two and three robot cases users were asked how much they preferred each interface (one user did not answer for the one and three-robot cases). The results, shown in Table 5.1, clearly show that people preferred the toy interface over the touch interface in the two and three robot case. This preference echoed in the written questionnaires and post-test interview as well. One user explained that the toys gave them a "sense that [they were] in contact with the robot," and seven participants wrote that they found it easier to visualize the robot position and orientation with the toy interface. One user reasoned that the toys provide more visual cues about the orientation and organization than the flat images used in the touch interface.

	1 robot	2 robot	3 robot
Toy	10	19	17
Touch	10	4	3
Neither	2	0	2

Table 5.1 - User-preferred interfaces for each robot case (numbers represent users)

	Question regarding robot use	toy	touch
2-robot	Easy to control both simul.	78%	57%
	Worked with both simul.	70%	43%
	Worked with one at a time.	35%	74%
	Used both hands simul.	61%	43%
3-robot	Easy to control all three simul.	74%	48%
	Worked with all three simul.	39%	26%
	Worked with one at a time.	61%	61%
	Used both hands simul.	70%	52%

Table 5.2 - Percentage of users that responded positively to questions about using both hands and controlling multiple robots simultaneously.

5.3.4 Touch

Users described the touch interface as being simpler due to less equipment and more precise and accurate due to the high resolution of the screen. Further, the touch was reported to be less intimidating because it was familiar and more similar to traditional PC interfaces. On the other hand, many people complained of the RNT scheme, with eleven users explicitly reporting that it was unintuitive to rotate the robot icon around the finger point. This is a property of RNT that users liked for ballistic movements but which caused problems for precise rotation of the robot once it was at the target location (this matches previous findings regarding RNT (Kruger et al., 2005)). RNT rotation moves the center of the object, requiring a final corrective translation. Instead, users recommended that it would be more intuitive for the robot icon to rotate around the center, "spinning like a plate."

Finally, with the three-robot case a few users complained of visual clutter - 3 icons for the real robots, 3 icons for the robot-controlling widget, lines connecting them and the green halos crowd the interface. One participant complained that "for the touch interface, you have six pictures (displayed on the table). It becomes confusing (when they overlap on top of each other)."

5.3.5 Toy

Users reported that the toys "were tactile and seemed more realistic" with their three-dimensional nature, with seven users explicitly noting that with the toy it was "a lot easier to visualize what was happening [remotely]" and to visualize the robot configuration. Further, it helped make it "easier to understand the rotation" and other robot state, enabling them to "focus on collision avoidance."

The primary complaint (mentioned by several users) is that the reflective markers for the tracking system get in the way of grasp, where occluding the markers can make the system lose track of the toys and cause erroneous robot movements. Users reported that the marker areas become no-hands zones that distract users from the natural grasp-intuitiveness of the toy.

5.3.6 Robot Movements

Users complained through comments and feedback that the robots often moved unexpectedly, despite the contrary evidence shown in Figure 5.5, saying that it was often difficult to visualize the path that the robot would take and that the "robots seemed to take slightly different paths (than the one [users] planned)." The primary reason cited behind this is that users expected the robots to copy or replay the movements given by the user, including sidesteps and exact paths, instead of moving directly toward a landmark target as the robots were programmed to do. This was explicitly described by ten of the users, and the problem was more prominent overall in the three-robot case and with the toy cases.

Another aspect of this was that the robots did not move consistently or in a straight line due to physical constraints and noise such as the robot mechanics and a somewhat uneven carpet. Because of this, robots sometimes had to correct their trajectory in mid-movement. Users further pointed out that our interfaces gave them no indication of the robot moving and rotation speed, or time to target location.

The robots have mechanical limitations and challenges with precise movements. As such, they sometimes had difficulties moving to the exact target location specified by the user, and are sometimes off by as much as 10 cm. When this happened it was very obvious and visible to the user and in the worst cases added considerable visual clutter.

With the toy interface, moving an object from one place to another was reported to be a trivial task by most participants. However, one participant said that "at times [she] forgot [she] was moving a robot and not only toys", such that she would "pick up the first one and put it [at the target location] and then disregard" the robot, eventually resulting in collisions." However, with the touch interface, the same participant said that "if [the control] is on the screen, [she] is more likely to pay attention to where [the robots] are."

5.3.7 Collisions

By far, the primary user complaint overall was that the robots often collided in the multi-robot cases, with 15 users bringing it up in their written comments as making them change their approach and increasing the effort required. Collisions were not dramatic (i.e., there were no loud noises or damaged robots), but it often took the user special effort to separate the robots as they would push against each other. This really annoyed a few users, and several stated that they expected the robots to be smart enough to avoid each other. As five participants explicitly pointed out, users have to learn each robot's movement characteristics in order to make an efficient path plan and avoid collisions.

5.3.8 Two-Handed Interaction and Multitasking

One aspect we looked at is how users utilize their hands in the experiment and if they use both at the same time. Table 5.2 summarizes our findings, which are echoed in the user comments, showing how users found the toy easier than touch in general for simultaneous hand use, and for the two-robot case the toys were used to work with both robots simultaneously rather than one at a time as they did with touch. In the three-robot case, however, users generally worked with one robot at a time for both the toy and touch interfaces.

Users reported that it was easier to operate robots simultaneously when the movement paths were similar and parallel rather than different and crossing, and more specifically they resorted to sequential movements when they felt that collisions were likely. Conversely, referring to the touch interface one user said: "whenever we use both the hands there are strong chances of [sic] robots getting collide with each other."

5.3.9 Complexity

We found a correlation between the number of robots and certain properties of the user responses. First, the conviction behind user response (how strongly they agree or disagree) decreased as the number of robots increased. Figure 5.6 shows the breakdown of how strongly users responded to four core questions asked throughout the experiment across the one, two, and three-robot cases, independent of the interface used, clearly outlining the trend to *weaken* their stance with the increasing number of robots. Further, the number of complaints (primarily regarding collisions) from the users in both the written questionnaires and during the experiment greatly increased as the number of robots increased. The trends of responses shown in Figure 5.5 suggest a general weakening of ease of use and control over the robot with the increased number of robots.

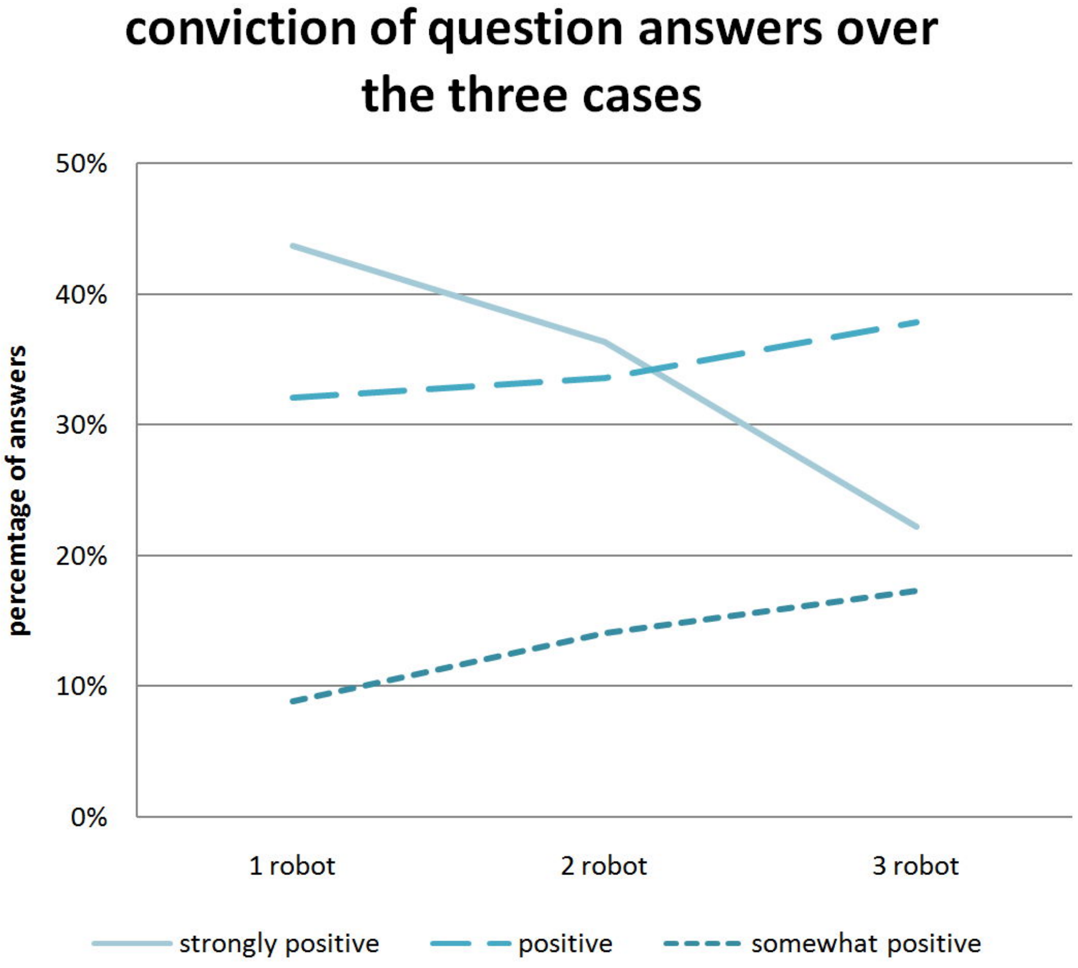


Figure 5.6 - The strength of user answers across the

5.3.10 Real Robots

In the post-test questionnaire, users were asked if they thought the experiment should have been done with a simulation instead of real robots. 15 of the 23 users stated that they felt that having real robots added value to the experiment. Reasons range from simple "the real thing is better" and "it is cool with real robots, more interesting than a simulation" to "real robots experience real problems. sims do not", "we trust the results more with real robots", "there was a real sense in knowing that real robots were colliding and that gave the situation more importance", and "real robots and the monitoring provided me with a better understanding of speeds and limitations that a simulation would have a hard time to capture."

5.3.11 TUI Evaluation

Here is the TUI evaluation we conducted based on Table 3.1 presented in Chapter 3:

TUI Benefits	Status
Physical I/O	Unsupported – Physical actions performed on the toys do not cause the robots to produce corresponding physical reactions.
Physical Feedback	Unsupported – The toys do not provide physical feedback to its holders. The robots' statuses are conveyed through visual feedback which is displayed on the tabletop computer.
I/O Unification	Supported – Users' action and perception spaces are unified on the tabletop computer's surface. Users received immediate visual feedback from the tabletop computer once toys are moved on the table surface.
Affordance	Supported – The color, shape and size of the toys directly reflects the color, shape and size of the robots. Participants can easily hold and move and rotate the toys.
Spatial Arrangement	Supported – Although different group formations of the robots do not infer or convey extra meanings to the participants in the task that we came up with, the system that we create easily afford the spatial arrangement

	property.
Spatial Mapping	Supported – The location and orientation of the toys are directly mapped to the location and orientation of the remote robots. Since the size of the table and the toys are similar to the experiment ground and the real robots, participants can understand the mapping between the toys' movements and the robots' reaction easily.

Table 5.3 – TUI Evaluation

5.4 Discussion

5.4.1 Task Completion Time

For the one-robot case, we observed that there is a significant difference between the Roomba and the AIBO in terms of task completion time. We believe this is mainly due to the mechanical preciseness of each robot. The Roomba moves on two wheels. It can move and rotate a lot more precisely than the four-legged AIBO. During the experiment, we constantly observed that the AIBO would "miss" (overshoot) its target location and turn around and walk back to correct itself. Thus, it usually takes the AIBO longer to move itself onto the target point.

For the two-robot case, we observed a significant difference between the two interaction methods. But for the three-robot case, there is no significant difference between the toy and touch interface. This phenomena may suggests that when a task is simple (like the one-robot case) or complicated (like the three-robot case) to perform, there is no real difference in terms of efficiency gain for both interfaces. In order to further improve the efficiency of such system, we may need to focus on simplifying high-level task planning (e.g. path planning) and collision avoidance.

It is hard to find solid evidence from the task completion time data to support any claim at this stage. This is because it is impossible to regulate the test administrator to move the target robots (shown in Figure 5.4) in a constant speed across all the test cases. Also, the participants were not limited to any route to make the robots to reach the target location. Thus, the task completion time may various significantly depending on the participants' controlling style. Therefore, further investigation is needed in order to draw a more solid conclusion.

5.4.2 Collisions and Cognitive Load

Collisions between robots were a large problem, slowing down the task, frustrating users, and increasing the concentration and effort necessary to complete the task. Given the importance that users gave this problem and the descriptions they gave in the written feedback, we feel confident in directly linking increase in collisions to the drop in user rating of ease-of-use and the resorting to only using one robot at a time in the three-robot case.

The data shows very clearly that the increasing concern with collisions was due to the users having more robots to worry about - more things to do and monitor at once puts higher demands on the user. It follows, then, that many of the collision-related complaints and problems are perhaps more accurately (and more simply) attributable to increased demand on the user, with collisions being another affect of this core problem. This agrees with Drury et al.'s human-robot interaction awareness taxonomy (Drury et al., 2003) and adds support for their claims regarding how the human-to-robot ratio affects interaction. What we found particularly surprising is how discernible this effect was in our experiment, where we only use three robots with very simple control mechanisms.

The number of robots is but one factor that influences user experience and usability. As the number of robots increases so does the demand on the user mental load, making it more difficult to compensate for interface limitations, which become more noticeable. This means that awareness and control problems will scale with the number of robots, and as such even seemingly minor interface flaws can become crippling.

The fact that a user reported paying more attention to a touch interface may suggest that although hiding low-level interface details from the users reduces their cognitive load, it can at the same time hinder their HRI awareness, and may lead them to forget certain important aspect of the task, possibly leading to undesirable incidents (such as collisions).

5.4.3 TUI and Touch

The very strong disparity between the results for the touch and TUI interfaces, and the fact that it solidified with more robots, is a strong indicator that our TUI interface was better suited to the task than our touch interface. Our data and findings frame a TUI vs touch set of results, but we must be careful with which conclusions we draw. User complaints with our touch implementation focused on the RNT technique, but had an overall effect on how *touch* was perceived. Applying our results to other touch interfaces needs to be done with care, and further experimentation will be necessary before drawing strong TUI vs touch-type conclusions.

5.4.4 Interface Design

User feedback directly outlined several problems with our interfaces. Both interfaces should be improved to afford the limitations and movement properties of the robots and the fact that the robots move in a straight line (and do not replay user input) to alleviate problems of the interface not being intuitive. Alternatively, we need to consider other interface styles, such as enabling users to specify either a path or a target. Further, our interface could improve problems of visual clutter (e. g., when the robot did not line up perfectly with the input), impairing user ability to concentrate on their task. This has further implications for the toy interface, as the inaccuracy damages the input-output unification: while the robot is supposed to be where the toy is, the error reminds the user of the separation, a fact they have to consciously compensate for.

5.4.5 User Experience and Emotion

The users strongly favoured the toy interface in most respects. Our results strongly link this success to core TUI concepts, as users explicitly and continually commented on the intuitive usability, the awareness gains, and the enjoyment they gained with the interface. This finding is quite significant and suggests that TUI interfaces should be explored in more depth for the remote control of robots.

Despite this, however, toy and touch interfaces were equally *efficient* in terms of the time required to complete the one-robot and three-robot tasks. We believe that this points to a deeper, but perhaps simple, dimension to our results. The users simply found the toys *fun* and *felt* connected to the robots when using them, which had a direct effect on how users felt about the usability of the interface. This is similar to how users defended the use of real robots due to the *cool* and novelty factor. These findings directly correspond to recent arguments for the incorporation of emotion and feeling into design, and HRI specifically (e. g., (Norman, 2004)).

5.4.6 Two Hands or One

The question of exactly when two-handed interaction is more effective is beyond the scope of our work, but in our experiments users resorted to one-handed interaction as things got complex, confusing, or difficult. This can be seen as another indicator of mental load, and a benefit of simpler interfaces - they may promote multi-hand interaction and the versatility that comes with it.

5.4.7 Implications

Here we distill our findings into a set of initial lessons and implications relevant for designing tabletop, touch, and TUIs for interaction with a remote group of robots.

- Users should not be expected to extrapolate the robot path, speed, and task just from the robot motions, but instead the interface should clearly indicate these properties to aid users in planning and interaction and to improve their HRI awareness.

- TUIs have a strong impact on user experience, regardless of particular efficiency gains, that can change how an interface is approached, perceived, used, and evaluated.
- Enabling users to specify complex, multi-part paths and commands relating to macro-scale robotic actions reduces their involvement and helps them cope with more robots in complex interaction scenarios.
- Users need to resort to lower-level control when the autonomy of the robot cannot solve a problem, such as a navigation complications or collisions. Good design should support this layer of detailed interaction as a backup option.
- Users may utilize both hands when interacting with a group of robots through tabletop, touch and TUIs. However, users may resort to single-hand interaction when they are faced with increasing cognitive load.
- Using actual robots (and letting the user know) changes the interaction experience in real ways that designers need to consider.

5.5 Conclusion

In this chapter, we have presented two novel interfaces and implementations for remotely interacting with multiple robots in real time using TUIs and touch. These interfaces support small groups of robots, using a tabletop computer as an interaction surface and provide detailed visual feedback on the robot location, state, and trajectory to enhance the user HRI task awareness. By conducting an empirical study of simple robot movement and group formation tasks, our analysis revealed several important relationships between the user experience and the properties of the interface used. We present the findings as a set of guidelines that researchers can use in their own interface design for remote robot control.

Chapter 6. Conclusion

In this thesis, we suggest, implement and test the concept of applying tangible user interfaces to human robot interaction. In particular, we motivate the use of TUIs as a beneficial HRI design approach (Chapter 3), and we present our designs, implementations and evaluations of two tangle user interfaces we have created for interacting with a single robot (Chapter 4), and a group of robots (Chapter 5). In this final chapter, we revisit our original research questions, summarize our thesis contributions, and conclude by discussing possible directions of future work relating to this research.

6.1 Research Questions Revisited

In Chapter 1, we raised the following four questions relating to apply TUIs to HRI:

1. *Why apply tangible user interface to HRI?* What kind of benefit do tangible user interfaces offer for tackling HRI-specific problems?
2. *How can we apply tangible user interfaces to allow a single operator to interact with a single robot?*
3. *How can we apply tangible user interfaces to allow a single operator to interact with a team of robots?*
4. *What are the advantage (or disadvantage) of using tangible user interfaces for completing HRI tasks compared to more orthodox, and to non-tangible user interfaces?*

6.2 Thesis Contributions

Following these research questions, our research efforts resulted in the following five main contributions:

1. *Proposing the Concept of applying TUIs to HRI*. In Chapter 3, we explained the design motivation behind applying TUIs to HRI. We explained our vision of how TUIs could help to alleviate and solve a set of HRI problems.

2. *Using TUI for single robot control*. In Chapter 4, we presented the design and implementation of a Nintendo Wii^{TM} controller based TUI for interacting with a Sony AIBO robot dog. To our knowledge, we are among the very first who demonstrated the usage of TUI for controlling a real robot. This Wii^{TM} controller based TUI allows people to control the locomotion and postures of an AIBO using arm and hand gestures.

3. *Comparative user study for single robot interaction*. In Chapter 4, we also described a keypad based interface that we created a as a reference system for evaluating the usability of the Wii^{TM} controller based TUI. We designed a comparative user study to allow people to test these two interfaces against each other by completing two robotic tasks. From the study, we found that the test participants completed both tasks faster and made fewer errors with the TUI. Also, most participants indicated that the TUI is easier to learn and use than the keypad interface.

4. *Using TUI for interacting with a team of robots*. In Chapter 5, we presented the design and implementation of a toy based TUI for interacting with a group of heterogeneous robots. To allow a single operator to control multiple robots, we used three plush toys tagged with reflective markers as handles for interacting with two AIBO robot dogs and one Roomba vacuum cleaner. The toys are place and tracked on a tabletop computer using a set of Vicon motion capture cameras, and are thus spatially mapped to the robots in real time. The tabletop computer is used to display the current status and location of the remote robots. To our knowledge, we are the first to suggest and implement many-to-many mapping of TUIs and a group of robots.

5. *Using touch based interface for interacting with a team of robots*. In Chapter 5, we also presented the design and implementation of a touch based interface for interacting with the same group of robots mentioned above. To our knowledge, we are the first who utilized a touch enabled tabletop computer for remotely controlling a group of robots. This touch interface allows people to control the location and orientation of remote robots by touching the corresponding icons displayed on the table. Also, two different robots can be controlled at the same time using two fingers.

6. *Comparative user study for multiple robot interaction*. In Chapter 5, we described a comparative user study between the TUI and the touch interface that we implemented. The test participants were asked to complete a simple robot navigation task using 1, 2 and 3 robots. From the study, we found that there is no real difference in terms of task completion time between the two interfaces. However, most people preferred the physical interaction, and have found the TUIs to be more enjoyable. From the data we gathered, we created a list of implications that may help other researchers to design effective interfaces for human-robot interaction using TUIs.

6.3 Future Work

The works that we have done in this thesis are just the first step in exploring the use of TUI in HRI. There are still a lot of works need to be done in order to fully realize the potentials of TUIs for interacting with robots. In this section, we describe possible future directions this research could take include improving the existing two TUIs that we have created for single and multiple robot interaction; further evaluation of the TUI (described in Chapter 5) using a more elaborate robotic task; and, exploration of alternate TUI designs to support more HRI tasks.

6.3.1 Improvements to the Wii^{TM} Controller Based TUI

We plan to improve the Wiimote/Nunchuk interaction technique to analyse continuous human arm and hand gestures to extend our abilities in controlling anthropomorphic and zoomorphic robots. We believe more elaborate TUIs would afford intuitive mapping for much more delicate HRI tasks. For instance, we can use an AIBO as a TUI to control the action of another AIBO. This would allow us to truly achieve a one-to-one mapping between the user's action and the robot's action. However, this degree of replication may not be ideal for repetitive tasks like walking. Thus, the degree of replication between the robot and the TUI has to be taken into account during the design stage.

Another approach to improve the usability of the Wiimote interface is to analyse the type of mistakes that participants made during the experiment. For the robot navigation task, we would like to identify the spots where the participants made mistakes on the obstacle course. Then, we can generalize the mistakes into different categories and make more in-depth analysis from there.

6.3.2 Improvements to the Toy Based TUI

With our studies being exploratory in nature, we believe our findings revealed only some of the basic lessons in using touch and TUIs for interaction with a robotic group, and we are planning to expand and improve on our experimentation.

We would like to explore an improved set of toys, ones that would contain more of the physical constraints of the robots. For example, we are planning to use toys with wheels that enforce the movement style and properties of the robots. As an extreme condition, we would like to test an interface based on a set of robotic TUIs that are identical to their coupled remote robotic team. Another, simpler approach we are considering is improvement in the visual feedback layer provided to the user (for both the touch and toy interfaces), such as a graphical template around the robot showing which directions it can move in. As technology advances, we may even be able to incorporate the deformable display technology that is envisioned in organic user interface (Vertegaal and Poupyrev, 2008) into TUIs to extend TUIs' ability in displaying digital information.

The current touch implementation brought to light interesting possibilities for improvement and we would like to explore how other touch techniques relate to our research problem, such as using touch gestures for moving the robot. Further, many of the physical properties of TUIs such as the three-dimensional nature or the natural collision detection can be ported to the touch interface, by restricting overlapping touch icons, or by using three-dimensional graphic visualizations rather than the current two-dimensional flat visualizations. We believe that improving our toy and touch interfaces will allow a more structured, and perhaps more conclusive, comparison between the two.

Another area that can be improved is the path-finding algorithm that the robots use for reaching the target location specified by the user. Currently, the robot would just walk in a straight path from its current location to the target location. The user cannot specific multiple way points for the robot to follow. The user has to wait until the robot to reach one location and then specify the next location. To solve this problem, we can record the way points specified by the user. Then, we can use algorithms like Hermite Curve Interpolation (Bartels et al., 1998; Catmull and Rom, 1974) or B-spline (Foley et al., 1990) function to calculate a smooth path among these points. This would make the path that the robot chooses to take smoother and looks more “natural” to the user.

The initial results suggest a correlation between one and two-handed use and the complexity of the task. We believe that this should be explored in more detail, both in terms of literature review and further experimentation focusing on the issue.

Mapping our touch and toy interaction approaches to more meaningful tasks will help us to validate our approach. We are considering experimenting with the robots in more valid tasks in lab setting. We are considering a group interface that will require the user to lead the robot through a simple spatial maze and will include collaborative tasks such as pulling and pushing objects. In the future, we believe our interfaces can easily scale to more meaningful robotic platforms and tasks, such as UAVs and USAR robots.

6.3.3 Thoughts on Future TUI Design for Zoomorphic Robots

Nature and our rich interaction with physical objects should inspire future research into designing and developing TUIs for HRI tasks. Specifically, in order to make TUIs more intuitive and accessible to non-expert users for controlling zoomorphic or anthropomorphic robots, we should consider utilizing the physical metaphors that are commonly observed in human-animal interaction for this propose. We believe that direct physical interaction techniques with robots will emerge from observing the extremely rich interaction techniques used by humans for domesticating animals, very similar to the reins metaphor that we used in the AIBO navigation task (Chapter 4). For example, we have seen collaborative hunting techniques using golden eagles, fishing techniques using cormorants, and the vast spectrum of existing interaction techniques between humans and dogs.

Animals are tamed and domesticated by humans for various proposes, examples range from providing labor, raising as food sources all the way up to forming intimate sociable relationships. In the case of training and utilizing animals as laborers, people use physical objects such as whip and rein to directly apply forces on the animals to reinforce their commands. These instruments, although very physical and aggressive in nature, provide instantaneous control and feedback for both the animal and the operator and, while ethically questionable, are very efficient. We believe this simple physical control mechanism can be very efficient for various collocated robotic interfaces. For instance, the BigDog robot (BigDog, 2008) build by Boston Dynamics is a carrier robot acts like a mule for transporting supplies on a battlefield. Such robots may need to deal with various interaction layers, some of them maybe as simple, physical and direct as a kick or whip.

6.4 Final Words

In this thesis, we have described the design, implementation, and evaluation of two tangible user interfaces that we have created for interacting with single and multiple robots. Studies have shown that people enjoyed using these TUIs for completing simple HRI tasks. Also, statistical results have revealed that TUIs can outperform traditional UI in terms of task completion time for certain HRI tasks. These findings suggest that there is a great potential in applying TUIs to HRI. We hope that the research described in this thesis defines a starting point in exploring the use of TUIs in HRI. We also hope it will motivate others to build upon and further extend our work to create better TUIs to allow humans to interact with robots more intuitively.

References

Adolph, K. E., Eppler, M. A., and Gibson, E. . (1993) **Development of perception of affordances**. In C. Rovee-Collier & L.P. Lipsitt (Eds.). Advances in Infancy Research. Norwood, NJ: Ablex. Vol. 8, 51-98.

Analog Devices. (2007). ADXL330. Retrieved June 13, 2008, from http://www.analog.com/UploadedFiles/Data_Sheets/ADXL330.pdf

Anasagasti, K., Bahrick, L. E., and Batista, L. C. (2002) **Perception of the Affordances of Everyday Objects by Human Infants**. International Society for Developmental Psychobiology, Orlando, FL.

Autodesk. **Maya**. Retrieved May 27, 2008, from http://usa.autodesk.com/adsk/servlet/index?id=7635018&siteID=123112

Balakrishnan, R., and Hinckley, K. (2000). **Symmetric bimanual interaction**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2000). ACM Press, 33-40.

Ballantyne, G., and Moll, F. (2003). **The da Vinci telerobotic surgical system: the virtual operative field and telepresence surgery**. Surgical Clinics of North America, Vol 83, Issue 6, pp 1293-1304.

Bartels, R. H., Beatty, J. C., and Barsky, B. A. (1998) **Hermite and cubic spline interpolation**. An Introduction to Splines for Use in Computer Graphics and Geometric Modelling. San Fransisco, CA. Morgan Kaufmann, 9 – 17.

Beaudouin-Lafon, M. (2000). **Instrumental interaction: an interaction model for designing post-WIMP user interfaces**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2000). ACM Press, 446-453.

Becker, M., Kefalea, E., Mael, E., Malsburg, C., Pagel, M., Triesch, J., Vorbruggen, J. C., Wurtz, R. P., and Zadel, S. (1999). **GripSee: A gesture-controlle drobot for oejct perception and manipulation**. Autonomous Robots. Vol 6, Number 2, 203-221.

Berkeley Robotics Laboratory. **BLEEX**. Retrieved May 27, 2008, from http://bleex.me.berkeley.edu/bleex.htm

BigDog. Retrieved May 28, 2008, from http://www.bostondynamics.com/content/sec.php?section=BigDog

Bluethmann, W., Ambrose, R., Diftler, M., Askew, S., Huber, E., Goza, M., Rehnmark, F., Lovchik, C. and Magruder, D. (2003). **Robonaut: A robot designed to work with humans in space**. Autonomous Robots (2003). 14(2-3): 179-197.

Breazeal, C. (1999). **Robot in society: Friend or appliance?**. In Agent 99 workshop on emontion-based agent architechtures. 18-26.

Breazeal, C. (2003). **Emotion and sociable humanoid robots.** International Journal of Human-Computer Studies. 59:119-155.

Cao, Y. U., Fukunaga, A. S., Kahng, A. B., and Meng, F. (1995). **Cooperative mobile robotics: Antecedents and directions**. In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 1995). Vol. 1, 226-234.

Catmull, E., and Rom, R. (1974) **A class of local interpolating splines**. Computer Aded Geometric Designs, Academic Press.

Dourish, P. (2001). **Where the action is: The foundations of embodied interaction**. MIT Press, Cambridge, MA, USA.

Drury, J, L., Riek, L., and Rackliffe, N. (2006). **A decomposition of UAV-related situation awareness**. In Proceddings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction (HRI ’06). ACM Press.pp 88-94.

Drury, J. L., Scholtz J., and Yanco, H. A. (2003). **Awareness in human-robot interactions**. In Proceedings of IEEE International Conference on Systems, Man and Cybernetics (IEEE SMC 2003). Vol. 1, 912-918.

Durlach, N. we., and Mavor, A. S. (1995). **Virtual reality: Scientific and technological challenges**. National Academies Press.

Encyclopedia Britannica. **Sundial**. Retrieved July 4, 2008, from http://www.britannica.com/EBchecked/topic/573826/sundial

Faisal, S., Cairns, P., and Craft, B. (2005). **Infoviz experience enhancement through mediated interaction**. In Proceedings of International Conference on Multimodal Interface (ICMI 2005). ACM Press, 3-9.

Ferketic, J., Goldblatt, L., Hodgson, E., Murray, S., Wichowski, R., Bradley, A., Fong, T. W., Evans, J., Chun, W., Stiles, R., Goodrich, M. A., Steinfeld, A., King, D, and Erkorkmaz, C. (2006). **Toward human-robot interface standards II: An examination of common elements in human-robot interaction across space enterprise**. American Institute of Aeronautics and Astronautics.

Fiorini, P., Ali, K., and Seraji, H. (1997). **Health care robotics: a progress report**. In Proceedings of IEEE International Conference on Robotics and Automation (IROS 1997). Vol.2, 1271-1276.

Fitzmaurice, G. W. and Buxton, W. (1997). **An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1997). ACM Press, 43-50.

Fitzmaurice, G. W., Ishii, H., and Buxton, W. (1995). **Bricks: Laying the foundations for graspable user interfaces**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1995). ACM Press, 317-324.

Fishkin, K. P. (2004). **A taxonomy for and analysis of tangible interfaces**. Personal and Ubiquitous Computing. Volu 8, issue 5, 347-358.

Foley, J. D., Dam, A., Feiner, S. K. and Hughes, J. F. (1990). **Computer graphics – Principles and practice**. Addison Wesley, 2nd edition.

Gibson, E. J. (1969). **Principles of perceptual learning and development**. New York: Appleton-Century Crofts.

Goodrich, M. and Olsen, D. (2003). **Seven principles of efficient human robot interaction**. In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2003). 3943–3948.

Greer, A. D., Newhook, P. M., and Sutherland, G. R. (2008) **Human-machine interface for robotic surgery and stereotaxy**. IEEE/ASMS Transactions on Mechatronics. Vol 13, No 3, pp 355-361.

Hasanuzzaman, Md., Zhang, T., Ampornaramveth, V., Kiatisevi, P., Shirai, Y., and Ueno, H. (2004). **Gesture based human-robot interaction using a frame based software platform**. In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2004). Vol.3, 2883-2888.

Hinrichs, U., Capendale, S., and Scott, S. D. (2006). **Evaluating the effects of fluid interface components on tabletop collaboration**. In Proceedings of Advanced Visual Interfaces (AVI 2006). ACM Press, 27-34.

Hong, J. W., Tan, X. N. (1989). **Calibrating a VPL DataGlove for teleoperating the Utah/MIT hand**. In Proceedings of IEEE International Conference on Robotics and Automation (IEEE ICRA 1989). Vol.3, 1752-1757.

Hornecker, E. and Buur, J. (2006). **Getting a grip on tangible interaction: A framework on physical space and social interaction**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2006). ACM Press, 437 – 446.

iRobot. **Roomba**. Retrieved May 6, 2008, from http://www.irobot.com/sp.cfm?pageid=122

iRobot. **PackBot**. Retrieved May 26, 2008, from http://www.irobot.com/sp.cfm?pageid=109

Ishii, H. and Ullmer, B. (1997) **Tangible bits: Towards seamless interfaces between people, bits and atoms**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 1997). ACM Press, 234-241.

Jacob, R. J. K., Ishii, H., Pangaro, G., and Patten, J. (2002). **A tangible interface for organizing information using a grid**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2002). ACM Press, 339-346.

Kaminka, G. A., and Elmaliach, Y. (2006). **Experiments with an ecological interface for monitoring tightly-coordinated robot teams**. In Proceedings of IEEE International Conference on Robotics and Automation (IEEE ICRA 2006). 200-205.

Kaminka, G. A. and Elmaliach, Y. (2006). **Single operator, multiple robots: Call-request handling in tight-coordination tasks**. In Proceedings of Distributed Autonomous Robotic Systems 7. Springer-Verlag. 103-113.

Kazerooni, H. (1990). **Human-robot interaction via the transfer of power and information signals**. IEEE Trans. on System and Cybernetics (1990). Vol.20, No. 2, 450-463.

Kruger, R., Carpendale, S., Scott, S. D. and A. Tang. (2005). **Fluid integration of rotation and translation**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2005). ACM Press, 601-610.

Keepon. Retrived May 26, 2008, from http://beatbots.org

Kiesler, S. and Hinds, P. (2004). **Introduction to this special issue on human-robot interation.** Human Computer Interaction Vol.19(1/2): 1-8.

Kortenkamp, D., Huber, E., and Bonasso, R. (1996). **Recognizing and interpreting gestures on a mobile robot**. In Proceedings of Association for the Advancement of Artificial Intelligence (AAAI 1996). Vol.2, 915-921.

Kozima, H., Nakagawa, C. and Yasuda, Y. (2007). **Children-robot interaction: A pilot study in autism therapy**. Progress in Brain Research, Vol 164, pp.385-400.

Kuzuoka, H., Oyama, S., and Yamazaki, K. (2000). **GestureMan: A mobile robot that embodies a remote instructor's actions**. In Proceedings of ACM conference on Computer Supported Cooperative Work (CSCW 2000). pp.155-162.

Mandryk, R. L., Scott, S. D., and Inkepn, K. M. (2002). **Display factors influencing co-located collaboration**. In Extended Abstracts of Computer supported cooperative work (CSCW 2002). ACM Press, 137-138.

Murta, A. D., and Roper, K. (2000). **Visualizing mobile robot sonar data using polygon set opeartions**. Visual Data Exploration and Analysis VII. 346-353.

Nabaztag. Nabaztag. Retrieved May 26, 2008, from http://www.nabaztag.com/en/index.html

Norman, D. A. (1988). **The psychology of everyday things**. BasicBooks.

Norman, D. A. (2004). **Emotional desing: why we love (or hate) everyday things**. BasicBooks.

Peek, B. (2007). Managed Library for Nintendo's Wiimote. Retrieved June 13, 2008, from http://blogs. sdn.com/coding4fun/archive/2007/03/14/1879033.aspx

Quigely, M., Goodrich, M., and Beard, R. (2004). **Semi-autonomous human-UAV interfaces for fixed-wing mini-UAVs**. In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004). Vol.3, 2457-2462.

Raffle, H., Parkes, A., and Ishii, H. (2004). **Topobo: A constructive assembly system with kinetic memory**. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2004). ACM Press, pp 647–654.

Richer, J., and Drury, J. L. (2006). **A video game-based framework for analyzing human-robot interaction: characterizing interface design in real-time interactive multimedia applications**. In Proceedings of the 1st ACM SIGCHI/SIGART Confernce on Human-robot Interaction (HRI 2006). ACM Press. pp 266-273.

Ricks, B., Nielsen, C. W., and Goodrich, M. A. (2004). **Ecological displays for orbot interaction: A new perspective**. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004). Vol 3, 2855-2860.

Rogalla, O., Ehrenmann, M., Zollner, R., Becher, R., and Dillmann, R. (2002). **Using gesture and speech control for commanding a robot assistant**. In Proceedings of IEEE International Workshop on Robot and Human Interactive Communication. 454-459.

Rogers, Y. and Lindley, S. (2004). **Collaborating around large interactive display: Which way is best to meet?** Interacting with Computers, 16(6): 1133-1152.

Scott, S. D., Grant, K. D., and Mandryk, R. L. (2003). **System guidelines for co-located collaborative work on a tabletop display**. In Proceedings of Computer supported cooperative work (CSCW 2003). ACM Press, 159-178.

Sharlin, E., Watson, B. A., Kitamura, Y., Kishino, F. and Itoh, Y. (2004). **On Tangible User Interfaces, Humans and Spatiality**. Personal and Ubiquitous Computing, Springer-Verlag.

Sony. **AIBO**. Retrieved May 26, 2008, from http://www.sony.jp/products/Consumer/aibo/

Sturman, D. J. and Zeltzer, D. (1994) **A survey of glove-based input**. IEEE Computer Graphics and Applications. Vol.14, Issue, 1, 30-39.

Tekkotsu framework, Retrived June 12, 2008, from http://www.cs.cmu.edu/~tekkotsu/

Theis, C., Iossifidis, we., and Steinhage, A. (2001). **Image processing methods for interactive robot control**. In Proceedings of 10th IEEE International Workshop on Robot and Human Interactive Communication. 424-429.

Uskarci, A., Alatan, A. A., Dindaroglu, M., and Ersak, A. (2003). **Robot mimicking: A visual approach for human machine interaction**. Computer and Information Science (ISCIS 2003), Springer-Verlag.

Vertegaal, R., and Poupyrev, we. (2008). **Organic User Interface**. Communications of the ACM. Vol 51, Issue 6.

Waldherr, S., Romero, R., and Thrun, S. (2000). **A gesture based interface for human-robot interaction**. Autonomous Robots. Vol.2, Number 2, 151-173

WiiLi. **Motion Analysis**. Retrieved June 13, 2008, from http://www.wiili.org/index.php/Motion_analysis

WiiLi. **Wiimote**. Retrieved June 13, 2008, from http://www.wiili.org/index.php/Wiimote#Motion_Sensor

Wikipedia. (2008) **Powered exoskeleton**. Retrieved May 20, 2008, from http://en.wikipedia.org/wiki/Powered_exoskeleton

Williams, K. W. (2004). **A summary of unmanned aircraft accident/incident data: Human factors implications**. Technical report DOF/FAA/AM-04/24. Washington, DC. Office of Aerospace Medicine, FAA.

Yanco, H. A. and Drury, J. L. (2004) **Classifying human-robot interaction: An updated taxonomy**. In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2004). Vol.3, 2841-2846.

Yanco, H. A., Drury, J. L., and Scholtz, J. (2004). **Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition**. Journal of Human-Robot Interaction. Vol. 19, 2004, 117-149.

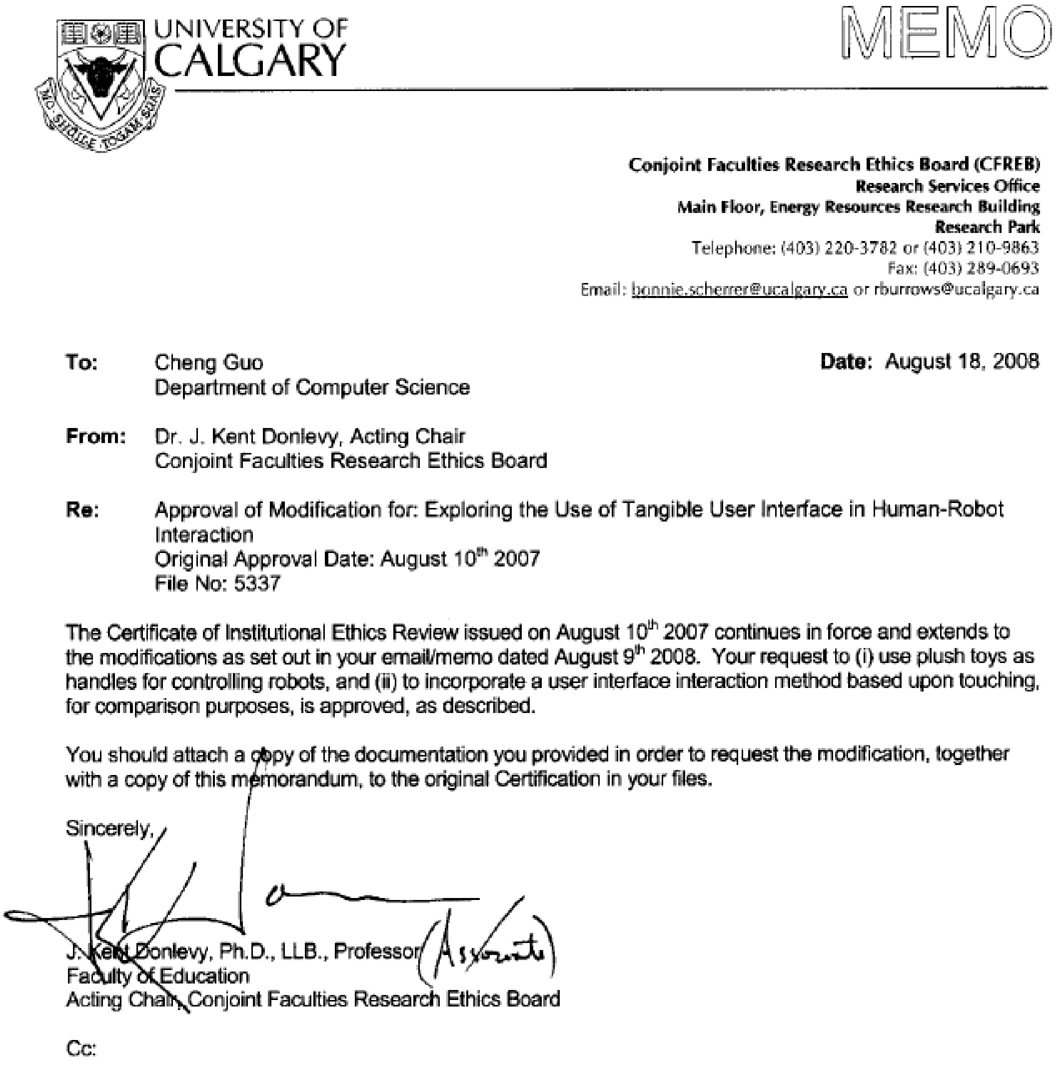
Young, J. E., Hawkins, R., Sharlin, E., and Igarashi, T. (2008). **Toward acceptable domestic robots**: Lessons learned from social psychology. University of Calgary, Computer Science Departmental Report. 2008-908-21.

Appendix A. Wiibot Study Materials

This appendix contains documentation related to the experiment procedure and evaluations described in Chapter 4. I have recruited 20 participants and the experiment was conduct in July 2007. The contents of this appendix are as follows:

1. ***Ethics Approval***: Ethics approval for this research was granted by Janice Dickin, Chair of the Conjoint Faculties Research Ethics Board at the University of Calgary.
2. ***Consent Form***: Study participants were required to read and sign this consent form prior to their interviews.
3. ***Experiment Protocol***: The test administrator followed this document to make sure that each experiment was conducted with the exact same procedure and condition.
4. ***Questionnaires***: Participants were asked to answer a list of questions before, during and after the experiment.

A.1 Ethics Approval



UNIVERSITY OF
CALGARY

MEMO

Conjoint Faculties Research Ethics Board (CFREB)
Research Services Office
Main Floor, Energy Resources Research Building
Research Park
Telephone: (403) 220-3782 or (403) 210-9863
Fax: (403) 289-0693
Email: bonnie.scherrer@ucalgary.ca or rburrows@ucalgary.ca

To: Cheng Guo
Department of Computer Science

Date: August 18, 2008

From: Dr. J. Kent Donlevy, Acting Chair
Conjoint Faculties Research Ethics Board

Re: Approval of Modification for: Exploring the Use of Tangible User Interface in Human-Robot Interaction
Original Approval Date: August 10th 2007
File No: 5337

The Certificate of Institutional Ethics Review issued on August 10th 2007 continues in force and extends to the modifications as set out in your email/memo dated August 9th 2008. Your request to (i) use plush toys as handles for controlling robots, and (ii) to incorporate a user interface interaction method based upon touching, for comparison purposes, is approved, as described.

You should attach a copy of the documentation you provided in order to request the modification, together with a copy of this memorandum, to the original Certification in your files.

Sincerely,

J. Kent Donlevy, Ph.D., LLB., Professor
Faculty of Education
Acting Chair, Conjoint Faculties Research Ethics Board

Cc:

A.2 Consent Form

Dr. Ehud Sharlin and Cheng Guo

Department of Computer Science

University of Calgary

2500 University Drive

Calgary, AB, CANADA T2N 1N4

Consent Form for Participants

Research Project: Exploratory study of Tangible User Interface in Human-Robot Interaction

Investigators: Dr. Ehud Sharlin and Cheng Guo

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Description of Research Project:

The purpose of this study is to explore the possibility of using Tangible User Interface in Human-Robot Interaction. You will be asked to control a Sony AIBO robot dog with two different controllers, the keyboard and Wiimote to complete two different set of tasks. You will be asked to repeat these tasks with different controllers. To start, we will teach you how to use a particular controller. After you have grasped the concept, you will be allowed to play with the AIBO until you are comfortable with the interaction techniques. Then, we will start the real experiment. The entire experiment will take 60 minutes. The experiment will be video taped and your comments about the experiment will be audio recorded.

Participation in this study will not put you at any risk or harm and is strictly voluntary. You choose to participate by playing the AIBO with two different controllers. You may choose to withdraw from the study at any time by simply not using the system any more. Any data collected to your withdrawal will still be available to the investigators for analysis.

Personally identifiable information will only be used in papers or presentations with your explicit permission. If we wish to use any personally identifiable information, we will contact you with the particulars of the information we wish to use, and you may decide whether or not you give us permission to use it. In this study, the personal information we will collect are your name, age and handedness which will be used only for identification purposes and grouping results. There are several options for you to consider if you decide to take part in this research. You can choose all, some or none of them. Please put a check mark on the corresponding line(s) to grant me your permission to:

I grant permission to be audio taped:
Yes: ___ No: ___
I grant permission to be videotaped:
Yes: ___ No: ___
I grant permission to have video or still images of me used in publications and/or presentations:
Yes: ___ No: ___
I grant permission to have comments of me used in publications and/or presentations:
Yes: ___ No: ___

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

At the conclusion of the study and its analysis, we will post any resulting papers that we have written about them. You can view these by asking the investigator or by accessing the website: http://grouplab.cpsc.ucalgary.ca/papers/index.html

Electronic data will be stored in a secure manner, such as in a computer secured with a password. Hardcopies of data will be stored in a locked cabinet/room with restricted access. Data will be kept for a minimum of three years and a maximum of 7 years. On disposal, electronic data will be erased and hardcopies will be shredded.

If you have further questions concerning matters related to this research, please contact:
Dr. Ehud Sharlin (403) 210-9499
ehud@cpsc.ucalgary.ca

If you have any concerns about the way you've been treated as a participant, please contact Bonnie Scherrer in the Research Services Office, University of Calgary at (403) 220-3782; email bonnie.scherrer@ucalgary.ca.

____

Participant's Signature Date

Investigator and/or Delegate’s Signature Date

Witness’ Signature Date

A copy of this consent form has been given to you to keep for your records and reference.

A.3 Experiment Protocol

Exploring the use of Tangible User Interface in Human-Robot Interaction – Protocol

<Remarks in brackets are directed for the administrator only>

1. Today is: **The experiment takes place in**:
2. Verify constant physical conditions:

Obstacle Course
Length: 262cm Width: 15.3cm Bright light conditions

Practice Trial 1
The cylinders are 26cm apart from each other.

Practice Trial 2
A twine is about 50 cm long which is pined and separated into three different sections.

4 Charged Batteries / 2 Wiimotes / 2 Nunchuks / 1 OQO / 1 Bluetooth Adapter

3. Introduction
"Hello, my name is Cheng. Today, we will perform an experiment involving human-robot collaboration. I'll briefly describe the concept of our project, talk about the prototype application that we have developed, and the procedure of the experiment."

"The goal of the experiment is to compare and contrast two different interaction techniques for controlling a Sony AIBO robot dog. The experiment consists of two parts, for the first part, you will be asked to navigate the AIBO through an obstacle course. For the second part, you will use the controllers to teach the AIBO to perform different postures. During the experiment, you will be observed and data will be collected for further analysis. The entire experiment will be video taped. Also, we will audio record some of the questionnaires that will be given to you during and after the experiment. You may quit the experiment at anytime if you don't want to continue."

"Would you like to participate in our study?"

4. Signing of the consent form
5. Participants are asked to complete a pre-study questionnaire

6. Participants will be trained for experiment Part 1
"Now, you will be trained to learn how to use the controller in your hand (Wiimote

/Keyboard) to navigate the AIBO dog. First of all, since we are controlling the AIBO through a wireless network, there is a half second network delay when you send a command to the AIBO. That means the AIBO will react a little bit slower than you would expect. Thus, when you want to navigate the AIBO, you should plan ahead and send commands to the AIBO in advance to compensate the network delay. Ok?"

<Answer any questions the participant may have>

"Now, I will introduce the basic controls to you."

<Introduce the Basic Controls: Walking Forward, Walking Forward & Turning Left, Walking Forward & Turning Right, Strafe Left, Strafe right, Rotate Left, Rotate Right & Stop. The participant will be given 2 minutes to get familiar with these basic movements >

"Ok, you have learned the basics, we can now move on to the next level. For this part of the training, you have to navigate the AIBO between two cylinders and walk in an "8" shape. While you are turning the AIBO around the cylinders, try to make a turn as tight as possible."

Figure A.1 - Practice Trail 1

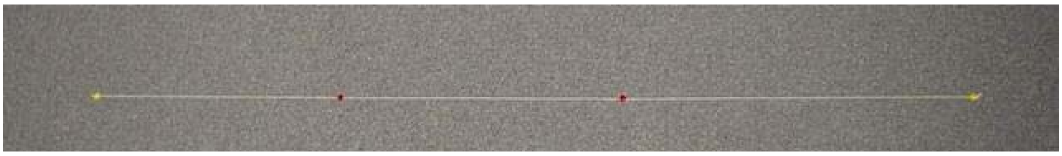
The cylinders are 26cm apart from each other.

< The participant will be given 2 minutes for this practice session >

"That was very good. Now, let's get into the last phase of our training session. In this section, I want you to combine walking, rotation and strafing together without stopping the AIBO during the transition. You start off at one end of the pined twine. Then, you walk the AIBO towards the other end of the line. When the AIBO reaches the first red pin, you rotate the AIBO to the right until its body is perpendicular to the twine. Then, you control the AIBO to strafe left to reach the second red pin. After you reach the second red pin, rotate the AIBO to the left to align the dog with the twine and then keep walking the AIBO towards the end of the twine."

Figure A.2 - Practice Trail 2

A twine is about 50 cm long which is pined and separated into three different sections.



< The participant will be given 2 minutes for this practice session.>

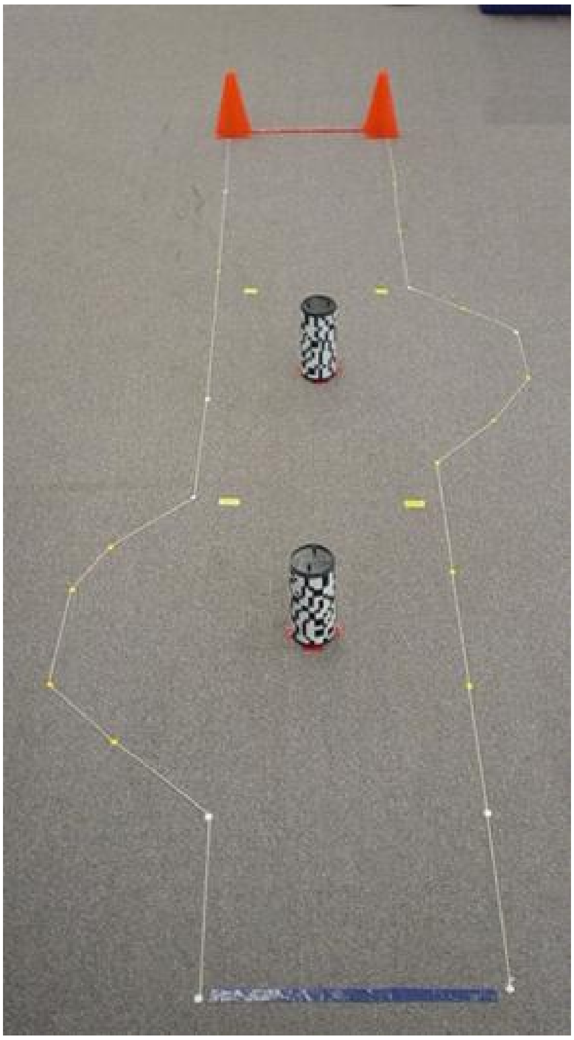
7. Participants perform the experimental conditions

“You have finished all of the practice Trials and now it is time to start the real experiment. The real experiment contains two difficulty levels. You will finish the easier level first followed by the harder level. I want you to guide the AIBO to run through the obstacle course as fast as possible without hitting any obstacles and try to maintain the AIBO inside the boundaries. You start behind the blue line. You finish the Trial when both of the back legs of the AIBO pass the red line. During the Trial, you will always stand behind the AIBO and follow it until the Trial ends. If you find the AIBO goes out of the boundary, you have to stop the AIBO immediately and guide it back to the Trial. Keep in mind that I will time the entire Trial so try to make as few mistakes as possible.”

<Explain how to walk through the easier obstacle course and then start the experiment.>

Figure A.3 - The Easier Obstacle Course

Length: 262cm Width: 15.3cm



“Very well done. It seems that the easier Trial is a piece of cake for you. Now, I will ask you to try a more challenging Trial. The rules for walking through this Trial are the same as the previous one. Try to not bump into the obstacles and make sure your AIBO stays inside of the boundary. Again, I will time you on this Trial.”

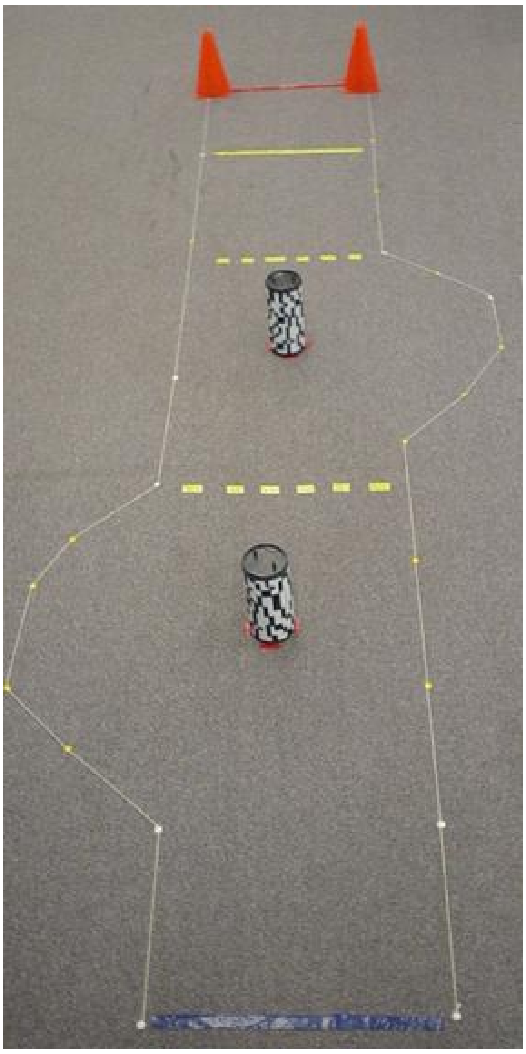
< Explain how to walk through the harder obstacle course and ask the participant to walk through the Trial without navigating the AIBO>

“If you forgot to perform a specific movement at certain locations of this obstacle course, I will have to move your AIBO back to the previous location and you have to perform the correct movement in order to pass through it. All of the time that is spent for correction will be count into the final finishing time."

<Start the experiment.>

Figure A.4 - The Harder Obstacle Course

Length: 262cm Width: 15.3cm



"That was challenging, wasn't it? But for a first time user, you did very well. Now, I'd like you answer the following questions and tell me what you think about this controller."

8. In-between study questionnaire.

<Use Audio Recorder to record the participant’s comments and thoughts about the controller>

“Great! Now the first part of the experiment is done. Let’s get started with the second part.

9. Participants will be trained for experiment Part 2

“In this part of the experiment, I am going to ask you to use this controller (Wiimote /Keyboard) to command the AIBO to perform two sets of postures with its forearms. Before we get started, I want you to run through the following Trials to get yourself familiar with the controller.”

Figure A.5 – Practice Gesture Combination 1

The participant will learn how to perform the above four postures from left to right.

“Here are four postures that I am going to teach you. The images you see here are reversed like mirrored images. But when you perform these postures, you don’t have to reverse them. If you see the AIBO raised the left arm (but from your perspective, it is the right arm), you raise your right arm. So basically, you perform the postures according to your own perspective. We have no intension to trick you in this experiment.”

“Remember that the lagging issue we had with the AIBO? We will have lagging problems in this part of the experiment too. This time, after you perform a posture, it will take the AIBO one to two seconds to perform it. Once the AIBO finishes a posture, it will bark to notice you that a posture has been performed. Sometimes, due to the lagging issue, the AIBO will not bark immediately after a posture is performed even though you can tell a gesture is indeed performed. During this circumstance, I want you to wait for the AIBO patiently. Only perform the next gesture after the AIBO barks.”

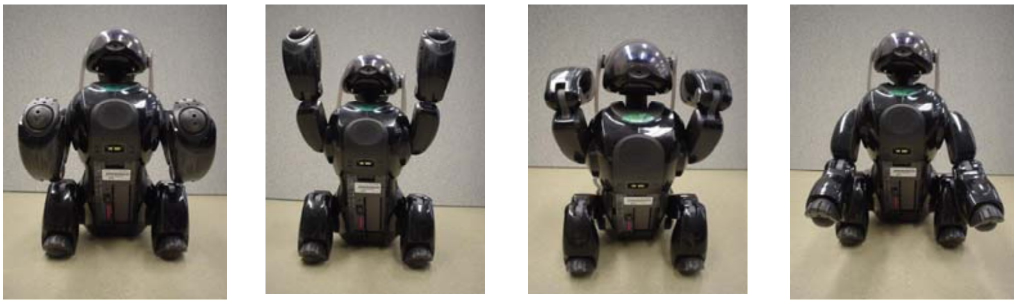
<Ask the participants if it is OK to put the Nunchuks on his/her arms>

<Teach the participant to perform the above postures with his/her right arm, then do the same thing with his/her left arm. This process will take 2 minutes.>

<* If the participant is using the wiimote interface, remind him/her that the wiimote has to always face downward when holding it. Also, when performing a posture, the participant should not use his/her wrists. The participant should keep his/her arms as close to his/her body as possible in order to get the best gesture recognition out of the system.>

"Well done. It seems that these postures are quite easy to perform for you. Now let's try some combined postures."

Figure A.6 – Practice Gesture Combination 2



The participant will learn how to perform the above four postures from left to right.

"In this practice Trial, I am going to teach you how to perform combined gestures. You can perform a gesture with both of your arms at the same time. Since, essentially, we are performing two separate gestures in total, we should hear two dog barks after the gestures are performed. After you hear two barks, you can keep continue on performing the next gesture."

<* If the participant is using the wiimote interface, remind him/her about the rules of using the wiimote. (Keep arms close to the body. Don't use wrist. Keep the wiimote facing downward when arms are down) >

<Teach the participant to perform the above postures. This process will take 2 minutes.>

10. Participants perform the experimental conditions

"Good job. Now it is time to do the real experiment. Just like the navigation experiment, I have prepared an easier and a harder Trial for you to test. Let's start with the easier Trial first: For this Trial, I will have six images to show you. The images will be displayed on my computer monitor. Once you see an image, you perform the posture shown on the image. After you are done, keep your posture. I will switch to the next image. The Trial ends when all six postures have been performed. If you performed a wrong posture, I want you to correct yourself immediately. During this experiment, I will time you. Do you have any questions?"

<Answer any questions the participant may have and then start the experiment>

Figure A.7 - The Easier Gesture Set

The participant will be asked to perform the above gestures from left to right in sequence.

<1 Minute Break>

“That was very good. Now let’s move onto the next level. The experiment procedure for this harder Trial is exactly the same as the easier one. Again, I will show you six images in sequence and you perform the postures accordingly. Although it may take you more time to perform these postures, please try to perform them as fast and correct as possible. Do you have any questions?”

<Answer any questions the participant may have and then start the experiment>

Figure A.8 - The Harder Gesture Set

The participant will be asked to perform the above gestures from left to right in sequence.

11. In-between study questionnaire

“That was a tough experiment, but you did very well. Here, I have couple of questions that I want to ask you about the controller.”

<Use Audio Recorder to record the participant’s comments and thoughts about the controller>

“Thank you for filling out the questionnaires. Let’s try another controller this time.”

12. Repeat Step 6 – 11 for the other controller

13. Post-study questionnaire & debriefing

“Thank you very much for your participation today! Now, you are very experienced with all of the controllers. I’d like to know your preferences among these controllers. Could you please answer these questions?”

<Use Audio Recorder to record the participant’s comments and thoughts about the controllers and the entire experiment>

14. Participants are paid

“Thank you very much for your participation today. Here is your payment. I am very appreciated for the effort that you have spent for the experiment.”

A.4 Questionnaires

Pre-Study Questionnaire

Age: _______ Handedness: Left-Handed or Right-Handed
Gender: Male ☐ Female ☐

1. How often do you use a computer keyboard?

1	2	3	4	5
Rarely	Yearly	Monthly	Weekly	Daily

2. How you ever played Nintendo Wii before?

Yes No

3. If you answered “Yes” to the question above, then how often do you play Nintendo Wii?

1	2	3	4	5
Rarely	Yearly	Monthly	Weekly	Daily

4. How often do you play computer games?

1	2	3	4	5
Rarely	Yearly	Monthly	Weekly	Daily

5. When you play computer games, how often do you use a keyboard to play?

1	2	3	4	5
Never	Rarely	Sometime	often	Very often

In Between-study Questionnaire

Part 1 – Navigation

1. With this controller, I can easily control the AIBO to where I wanted.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

2. I found the controlling scheme is easy to learn.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

3. I found the controller is difficult to use when rotating the AIBO.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

4. I found the controller is difficult to use when perform sidestepping.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

5.I found the controller is difficult to use when turning the AIBO while walking forward.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

6. I had difficulty remembering how to perform certain movements with this technique.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

7. Please add any other comments and/or thoughts about this controller.

Part 2 – Posture

1. I found this controlling technique is easy to learn.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

2. I found this controlling technique is easy to use.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

3. With this technique, the AIBO reacted as I expected it to.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

4. I had difficulty controlling both of the AIBO’s arms at same time using this technique.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

5. I had difficulty remembering how to perform certain gestures with this technique.

1	2	3	4	5	6	7
Strongly Disagree	Disagree	Weakly Disagree	Neutral	Weakly Agree	Agree	Strongly Agree

6. Please add any other comments and/or thoughts about this controller.

Post-study Questionnaire

1. For experiment Part 1 (Navigation), overall, which one of the controllers do you prefer to use? (Please circle one of the options below)

Wiimote Keyboard No preference

2. For experiment Part 2 (Posture), overall, which one of the controllers do you prefer to use? (Please circle one of the options below)

Wiimote Keyboard No preference

3. Which controller do you feel more natural to use for the Navigation task? (Please circle one of the options below)

Wiimote Keyboard No preference

4. Which controller do you feel more natural to use for the Posture task? (Please circle one of the options below)

Wiimote Keyboard No preference

5. Please add any other comments and/or thoughts about the controllers and the experiment.

Appendix B. Toy and Touch Study Materials

This appendix contains documentation related to the experiment procedure and evaluations described in Chapter 5. I have recruited 23 participants and the experiment was conduct in August 2008. The contents of this appendix are as follows:

1. ***Consent Form***: Study participants were required to read and sign this consent form prior to their interviews.
2. ***Experiment Protocol***: The test administrator followed this document to make sure that each experiment was conducted with the exact same procedure and condition.
3. ***Questionnaires***: Participants were asked to answer a list of questions before, during and after the experiment.
4. ***Implementation Detail***: This section explains the implementation detail of the project explained in Chapter 5.

Note: The ethics approval form obtained for this experiment is part of the ethics approval included in Appendix A.1

B.1 Consent Form

Dr. Ehud Sharlin and Cheng Guo

Department of Computer Science

University of Calgary

2500 University Drive

Calgary, AB, CANADA T2N 1N4

Consent Form for Participants

Research Project: RICON User Study

Investigators: Dr. Ehud Sharlin, Cheng Guo and Jim Young

This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Note: The University of Calgary Conjoint Faculties Research Ethics Board has approved this research study.

Description of Research Project:

The purpose of this study is to explore the possibility of using Tangible User Interface and Touch Interface for remotely controlling multiple robots. The entire study is divided into three parts. You will be asked to control one, two and three robots to follow target points in each part of the study. Before each study, we will teach you how to use each interface and let you practice first. After you have grasped the concept, we will start the real experiment. The entire experiment will take 60 minutes. The experiment will be video taped and your comments about the experiment will be audio recorded.

Participation in this study will not put you at any risk or harm and is strictly voluntary. You choose to participate by playing the AIBO with two different controllers. You may choose to withdraw from the study at any time by simply not using the system any more. Any data collected to your withdrawal will still be available to the investigators for analysis. Personally identifiable information will only be used in papers or presentations with your explicit permission. If we wish to use any personally identifiable information, we will contact you with the particulars of the information we wish to use, and you may decide whether or not you give us permission to use it. In this study, the personal information we will collect are your name, age and handedness which will be used only for identification purposes and grouping results. There are several options for you to consider if you decide to take part in this research. You can choose all, some or none of them. Please note that in any case we will not expose your name or identity. However, if you grant us permission, we may use your picture during interaction in academic publications/presentations about this research. Please put a check mark on the corresponding line(s) to grant me your permission to:

I grant permission to be audio taped:
Yes: ___ No: ___
I grant permission to be videotaped:
Yes: ___ No: ___
I grant permission to have quotations from my comments answers that are recorded during the study to be used in publications and/or presentations (note that your identity will never be associated with the quotations):
Yes: ___ No: ___
I grant permission to have video or still images of me used in publications and/or presentations:
Yes: ___ No: ___

If researchers wish to include information that may identify me, such as my picture or video, in reports of the data, I prefer the researchers to re-contact me for permission:
Yes: ___ No: ___ If Yes, Please leave your contact information:
____

____
___

Your signature on this form indicates that you have understood to your satisfaction the information regarding participation in the research project and agree to participate as a subject. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities. You are free to withdraw from the study at any time. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

At the conclusion of the study and its analysis, we will post any resulting papers that we have written about them. You can view these by asking the investigator or by accessing the website: http://grouplab.cpsc.ucalgary.ca/papers/index.html

Electronic data will be stored in a secure manner, such as in a computer secured with a password. Hardcopies of data will be stored in a locked cabinet/room located at the University of Calgary Interactions Laboratory (Math Science building, room 680) with restricted access. Data will be kept for a minimum of three years and a maximum of 7 years. On disposal, electronic data will be erased and hardcopies will be shredded.

If you have further questions concerning matters related to this research, please contact:
Dr. Ehud Sharlin (403) 210-9499
ehud@cpsc.ucalgary.ca

If you have any concerns about the way you've been treated as a participant, please contact Bonnie Scherrer in the Research Services Office, University of Calgary at (403) 220-3782; email bonnie.scherrer@ucalgary.ca.

Participant’s Signature	Date
Investigator and/or Delegate’s Signature	Date

A copy of this consent form has been given to you to keep for your records and reference.

B.2 Experiment Protocol

Ricon Experiment Protocol

<Remarks in brackets are directed for the administrator only>

1. **Introduction**

"Hello, my name is Cheng. Today, we will perform an experiment involving remote robot control. I'll briefly describe the concept of our project, talk about the prototype application that we have developed, and the procedure of the experiment."

"The goal of the experiment is to compare and contrast two different interaction techniques for controlling a group of robots. The two interaction techniques are Toy interface <show the toys to the participant> and Touch interface <briefly explain what it is>."
hmm. Make sure to introduce properly the table and how it works.
"The experiment consists of three parts, for the first part, you will be asked to navigate a single robot by following target points on the table. For the second and third part, you will control two and three robots respectively to complete the same task. During the experiment, you will be observed and data will be collected for further analysis. The entire experiment will be video taped. Also, we will audio record some of the questionnaires that will be given to you during and after the experiment. You may quit the experiment at anytime if you don't want to continue." Add a note here about privacy, anonymous data.

"Would you like to participate in our study?"

2. **Signing the consent form (sign, not sing)**
3. **Participants are asked to complete the pre-study questionnaire**
4. **Training for experiment part 1**

"Now, I am going to show you how to use the Toy/Touch interface for controlling a Roomba vacuum cleaner."

<Demonstrating one of the interfaces depending on the order>
<1) Tell the user that the robot's movement is imprecise>
<2) Tell the user that the green circle indicates that the robot has reached the target>

"Try it yourself. Once you feel comfortable with this interface, please tell me and we will start the real experiment. Please ask if you have any questions."

5. Experiment Part 1 Start

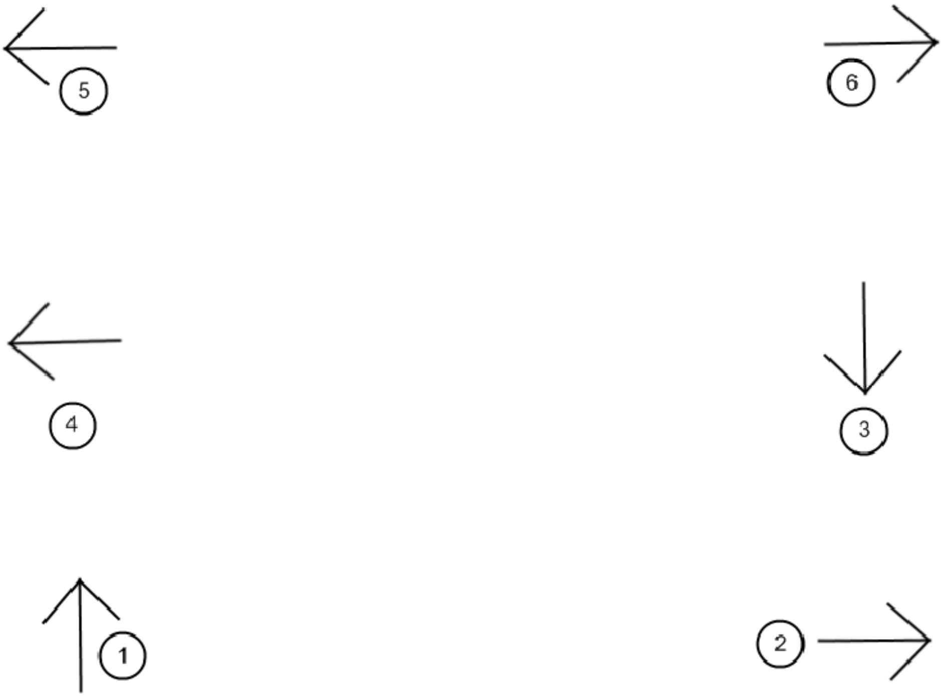


Figure B.1 – Task 1

<The arrows indicate the orientation of the robot. The number inside the circle indicates the order of steps. >

<Test administrator resets the robot (**Roomba**) at target 1>

“Now, I am going to use these images <Images of the actual robot printed on a piece of paper> to indicate the next target location and orientation of the robot. All you need to do is to make the robot to move to this location and align itself correctly. Once the robot stops (indicated by the green circle) then I will show you the next target location. We will repeat this process until I tell you the experiment is done.”

“You have probably noticed that due to the imprecise movement of the robot, it does not move onto the point where you want it to be. Instead, it will just somewhere close to it. When we run the experiment, you don’t have to worry about this problem. All you need to do is to move the toy/icon (depending on the interface) onto the location that I indicated on the white board. Any questions”

<Start the experiment> <hand-time how long they take>

“Great, now I want you to try another robot <**AIBO**> with the same task. I will let you try out this robot first and then we will repeat the previous experiment.”

<Repeat previous experiment>
Do a questionnaire before changing interfaces. We have one for touch and one for toy.l
<Change the Interface and repeat the task again>

6. In-between study questionnaire.
“Please fill out the questionnaire.”

7. Training for experiment part 2
“Now, we are going to start the second part of our experiment. For this part, I am going to ask you to control two robots on the table.” **–** do the same thing, waypoints

<Use **Roomba & AIBO** or **AIBO & AIBO** depending on the order>

“Please try out the interface and we will start the experiment.”

8. Experiment Part 2 Start

“The experiment procedure is going to the same as the previous experiment. I will show you the target point of both robots on the white board. You follow the waypoints until the experiment ends. When both robots stop on the target location, I will reveal the next location. Any questions? If no, then lets start.”

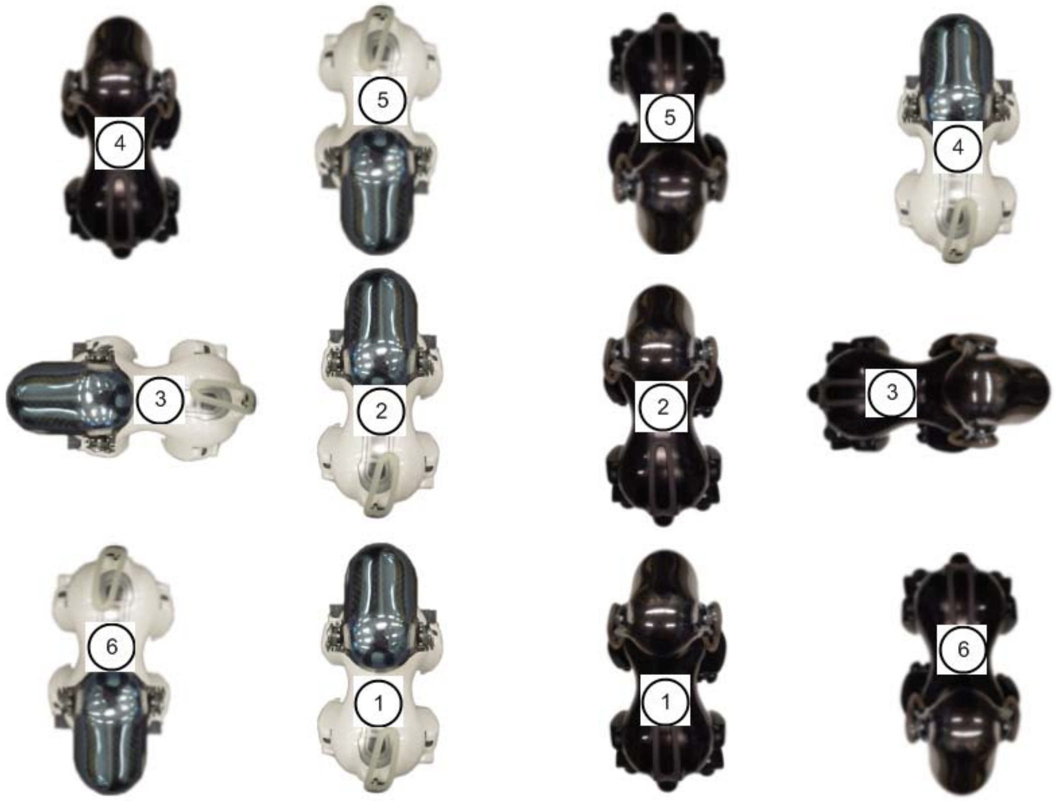


Figure B.2 – Task 2

<Change the robot set and allows the participant to practice until he/she is comfortable>

<Repeat the same task again>
<Change the Interface and repeat the task again>

9. In-between study questionnaire.

“Now we are done with the second part of the experiment. Please fill out the questionnaires.”

10. Training for experiment part 3

“Let’s start the last part of the experiment. This time, I am going to give you three robots for you to control. Please try out the interface and see if you have any questions.”

<Let the user practice with 3 robots>

“Just like the previous two experiments, I will show you the target location of each robot, you make them to move to their locations. Any questions?”

Figure B.3 – Task 3 Formation 1

<First formation>

<The green rectangle marks the start position. The red rectangle marks the target position>

Figure B.4 – Task 3 Formation 2

<Second formation>
<The participant has to change from the first formation to the second formation>

Figure B.5 – Task 3 Formation 3

<Third formation>
<The participant has to change from the second formation to the third formation>

Figure B.6 – Task 3 Formation 4

<Fourth formation>
<The participant has to change from the third formation to the fourth formation>

Figure B.7 – Task 3 Formation 5

<Fifth formation>
<The participant has to change from the fourth formation to the fifth formation>

<Change the Interface and repeat the task again>

11. In-between study questionnaire.
"Please fill out the questionnaires."

12. Post-study questionnaire & debriefing
"Thank you very much for your participation today. Now, you have done
all of the experiment. We'd like to know your overall feeling about this experiment."
<Ask the participant to fill out the post-study questionnaires>
Make sure to include interview 1-on-1 time in this

13. Pay the participant

B.3 Questionnaires

Pre-Study Questionnaire

Age: ______ Handedness: Left-Handed or Right-Handed or Ambidextrous
Gender: Male ☐ Female ☐

1. How familiar are you with touch-screen interfaces?

1	2	3	4	5
Never Seen Before	Never Used Before	Somewhat Familiar	Very Familiar	Expert

2. Do you have experience with remote robot control?

Yes No

If "Yes", then how familiar are you with robot remote controlling interface?

1	2	3	4	5
Only Used Once	Used a few Times before	Somewhat Familiar	Very Familiar	Expert

3. Do you consider yourself to be good at working with your hands? For example, good at woodworking or sewing?

Yes No

Do you consider yourself to have good hand-eye coordination?

Yes No

4. How often do you play video games?

1	2	3	4	5
Never	Yearly	Monthly	Weekly	Daily

One Robot Remote Control – TOY INTERFACE

To what extent do you agree / disagree with the following statements?
(if you feel there is no difference between the Roomba and AIBO, just fill out one set)

	Roomba							AIBO						
With the toy interface…	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…it was easy to move the robot to the target location (this does not include the orientation of the robot). **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was difficult to rotate the robot in the way that I wanted. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I had precise control over the robot movement.	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Comments:															
…the robots generally did not react as expected. **Comments**:	1	2	3	4	5	6	7		1	2	3	4	5	6	7

Any comments on the difference between controlling the two different robots?

Any additional thoughts or comments?

One Robot Remote Control – TOUCH INTERFACE

To what extent do you agree / disagree with the following statements?
(if you feel there is no difference between the Roomba and AIBO, just fill out one set)

	Roomba							**AIBO**						
With the touch interface…	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…it was easy to move the robot to the target location (this does not include the orientation of the robot). **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was difficult to rotate the robot in the way that I wanted. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I had precise control over the robot movement.	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Comments:							
…the robots generally did not react as expected. **Comments**:	1	2	3	4	5	6	7

1	2	3	4	5	6	7

Any additional thoughts or comments?

Any comments on the differences between controlling the two different robots?

(1 robot)	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
Overall, I preferred the touch interface **Comments**:	1	2	3	4	5	6	7
Overall, I preferred the toy interface. **Comments**:	1	2	3	4	5	6	7

Two Robots Remote Control – TOY INTERFACE

To what extent do you agree / disagree with the following statements?
(if you feel there is no difference between the robot configurations, just fill out one set)

	AIBO & AIBO							Roobma & AIBO						
With the toy interface…	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…it was easy to move the robot to the target location (this does not include the orientation of the robot). **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was difficult to rotate the robot in the way that I wanted. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I had precise control over the robot movement.	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Comments:														
…the robots generally did not react as expected. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was confusing to monitor both robots at the same time. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was easy to control the two robots at the same time **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I worked with both robots at the same time, operating them simultaneously **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7

…I worked with one robot at a time, operating them sequentially. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I often used both of my hands at the same time. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Any comments on the difference between the two robot configurations?
Any additional thoughts or comments?

Two Robots Remote Control – TOUCH INTERFACE

To what extent do you agree / disagree with the following statements?
(if you feel there is no difference between the robot configurations, just fill out one set)

	AIBO & AIBO							**Roobma & AIBO**						
With the touch interface…	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…it was easy to move the robot to the target location (this does not include the orientation of the robot). **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was difficult to rotate the robot in the way that I wanted. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I had precise control over the robot movement.	1	2	3	4	5	6	7	1	2	3	4	5	6	7

Comments:														
…the robots generally did not react as expected. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was confusing to monitor both robots at the same time. **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…it was easy to control the two robots at the same time **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7
…I worked with both robots at the same time, operating them simultaneously **Comments**:	1	2	3	4	5	6	7	1	2	3	4	5	6	7

…I worked with one robot at a time, operating them sequentially. **Comments**:	1	2	3	4	5	6	7		1	2	3	4	5	6	7
…I often used both of my hands at the same time. **Comments**:	1	2	3	4	5	6	7		1	2	3	4	5	6	7

Any comments on the difference between the two robot configurations?

Any additional thoughts or comments?

(2 robot)	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
Overall, for two robots I preferred the touch interface **Comments**:	1	2	3	4	5	6	7
Overall, for two robots I preferred the toy interface **Comments**:	1	2	3	4	5	6	7

Were there any particular changes or differences that you encountered with two robots that you did not find with the one robot?

Three Robots Remote Control – TOY INTERFACE

To what extent do you agree / disagree with the following statements?

With the toy interface…	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…it was easy to move the robot to the target location (this does not include the orientation of the robot). **Comments**:	1	2	3	4	5	6	7
…it was difficult to rotate the robot in the way that I wanted. **Comments**:	1	2	3	4	5	6	7
…I had precise control over the robot movement. **Comments**:	1	2	3	4	5	6	7
…the robots generally did not react as expected. **Comments**:	1	2	3	4	5	6	7
…it was easy to form the group formations. **Comments**:	1	2	3	4	5	6	7

…it was confusing to monitor all three robots at the same time. **Comments**:	1	2	3	4	5	6	7
…it was easy to control the three robots at the same time. **Comments**:	1	2	3	4	5	6	7
…I worked with all three robots at the same time, operating them simultaneously. **Comments**:	1	2	3	4	5	6	7
…I worked with one robot at a time, operating them sequentially.. **Comments**:	1	2	3	4	5	6	7
…I often used both of my hands at the same time. **Comments**:	1	2	3	4	5	6	7

Any additional comments or thoughts?

Three Robots Remote Control – TOUCH INTERFACE

To what extent do you agree / disagree with the following statements?

With the touch interface…	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…it was easy to move the robot to the target location (this does not include the orientation of the robot). **Comments**:	1	2	3	4	5	6	7
…it was difficult to rotate the robot in the way that I wanted. **Comments**:	1	2	3	4	5	6	7
…I had precise control over the robot movement. **Comments**:	1	2	3	4	5	6	7
…the robots generally did not react as expected. **Comments**:	1	2	3	4	5	6	7
…it was easy to form the group formations. **Comments**:	1	2	3	4	5	6	7

…it was confusing to monitor all three robots at the same time. **Comments**:	1	2	3	4	5	6	7
…it was easy to control the three robots at the same time **Comments**:	1	2	3	4	5	6	7
…I worked with all three robots at the same time, operating them simultaneously. **Comments**:	1	2	3	4	5	6	7
…I worked with one robot at a time, operating them sequentially. **Comments**:	1	2	3	4	5	6	7
…I often used both of my hands at the same time. **Comments**:	1	2	3	4	5	6	7

Any additional comments or thoughts?

(3 robot)	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
…overall, for three robots I preferred the touch interface. **Comments**:	1	2	3	4	5	6	7
…overall, for three robots I preferred the toy interface. **Comments**:	1	2	3	4	5	6	7

Were there any changes or differences with the three robot case that you did not notice or find with the one and two robot cases?

Post-Study Questionnaire

To what extent do you agree / disagree with the following statements?

	strongly disagree	disagree	somewhat disagree	no opinion	somewhat agree	agree	strongly agree
I found the graphical feedback on the table easy to understand. **Comments**:	1	2	3	4	5	6	7
The graphical feedback on the table was unnecessary. **Comments**:	1	2	3	4	5	6	7

Please describe the benefits that you noticed, if any, of the toy interface.

Please describe the benefits that you noticed, if any, of the touch interface.

Please describe the problems that you noticed, if any, of the toy interface.

Please describe the problems that you noticed, if any, of the touch interface.

Would you have rather done this experiment on a standard desktop PC? Why or why not?

You were controlling real robots. Instead, we could have done this with a simulation. Would this have been better? Why or why not?

Where else – besides robot control – could you imagine using the interfaces you used today?

B.4 Implementation Detail

We used two sets of Vicon cameras to capture the location of the toys on the tabletop computer and the location of the real robots at a remote place. There were seven cameras used for monitoring the tabletop computer area and four cameras for the robot area. Both sets of Vicon cameras are connected to their servers (called Nexus) for image analysis purpose. In order to combine the information gathered by both sets of Vicons and display the location information of the real robots on the tabletop computer, the Nexuses are connected to the tabletop computer through two separate Ethernet connections.

The set of Vicon cameras around the tabletop computer is used to track the location and orientation of the toys. By default, the cameras are running at 60 frames per second (fps). It is fast enough to capture the toys' movements. The toys' location and orientation on the table is used as the target location and orientation of the real robots.

The Vicon cameras used around the robot area track the location and orientation of the robots. This information is sent to the tabletop computer and is displayed on the table surface. Since the cameras are running at 60 fps and directly connected to the tabletop computer, participants receive real-time update about the robots' status on the tabletop surface.

The tabletop computer acts like a mediator in between the toys and the robots. It uses the information gathered from the toys to command the robots to react (through WiFi and Bluetooth). Also, it uses the information gathered form the robots to inform the participants about the current robots' status.

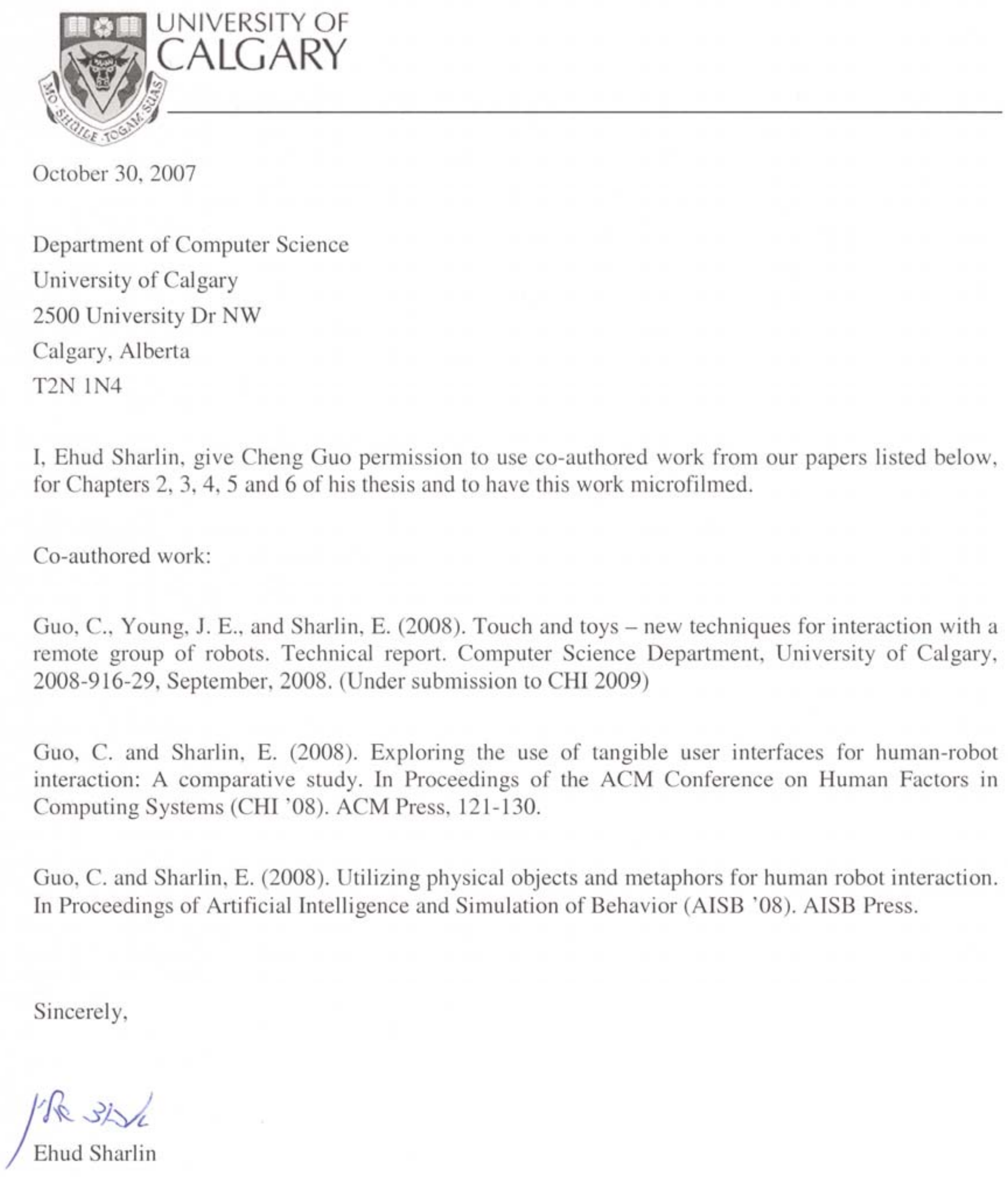
In order to distinguish among the toys and robots, we attached different number of reflective markers on the toys and robots. A “head” marker is attached to the head of each toy and robot. We calculate the centre of each toy and robot by averaging the positions of all the markers attached on them. By constructing a vector from the objects' centre to their “head” markers, we are able to find the orientation of the objects.

B.4.1 Path Finding Algorithm

We used a very simple path finding algorithm to control the robots to reach target locations specified by the participants. The path finding algorithm can be broken down into three steps: first, the robot would rotate itself to face toward the target location. Then, the robot would approach the target location in various speeds depending on the distance between the robot and the target location. If the robot is far away from the target location, it will walk or drive in a faster speed. As it gets close to the target location, it will slow down to prevent “overshoot”. If the robot deviates from its original path, it will stop and rotate to face toward the target location and start approaching again. Once the robot arrives on the target location, it would rotate itself to align with the direction that the toy is facing toward to.

Appendix C. Co-Author Permissions

In this appendix, I include permissions from my collaborators to use co-authored work from our papers in my thesis.



UNIVERSITY OF
CALGARY

October 30, 2007

Department of Computer Science
University of Calgary
2500 University Dr NW
Calgary, Alberta
T2N 1N4

I, Ehud Sharlin, give Cheng Guo permission to use co-authored work from our papers listed below, for Chapters 2, 3, 4, 5 and 6 of his thesis and to have this work microfilmed.

Co-authored work:

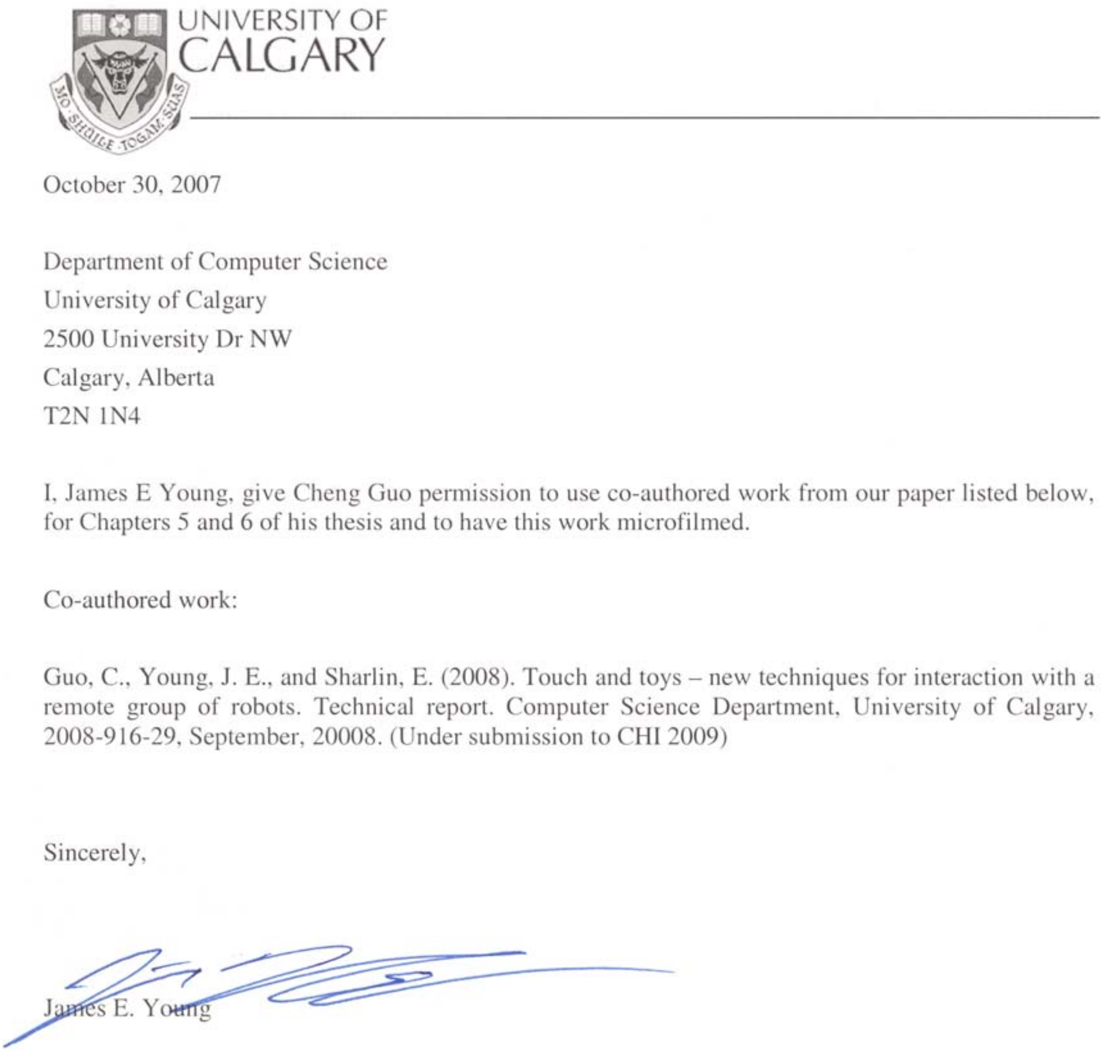
Guo, C., Young, J. E., and Sharlin, E. (2008). Touch and toys – new techniques for interaction with a remote group of robots. Technical report. Computer Science Department, University of Calgary, 2008-916-29, September, 2008. (Under submission to CHI 2009)

Guo, C. and Sharlin, E. (2008). Exploring the use of tangible user interfaces for human-robot interaction: A comparative study. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '08). ACM Press, 121-130.

Guo, C. and Sharlin, E. (2008). Utilizing physical objects and metaphors for human robot interaction. In Proceedings of Artificial Intelligence and Simulation of Behavior (AISB '08). AISB Press.

Sincerely,

Ehud Sharlin



UNIVERSITY OF
CALGARY

October 30, 2007

Department of Computer Science
University of Calgary
2500 University Dr NW
Calgary, Alberta
T2N 1N4

I, James E Young, give Cheng Guo permission to use co-authored work from our paper listed below, for Chapters 5 and 6 of his thesis and to have this work microfilmed.

Co-authored work:

Guo, C., Young, J. E., and Sharlin, E. (2008). Touch and toys – new techniques for interaction with a remote group of robots. Technical report. Computer Science Department, University of Calgary, 2008-916-29, September, 20008. (Under submission to CHI 2009)

Sincerely,

James E. Young