# Utilizing Physical Objects and Metaphors for Human Robot Interaction

**Cheng Guo**
University of Calgary
2500 University Drive NW
Calgary, AB, Canada
1.403.210.9404
cheguo@cpsc.ucalgary.ca

**Ehud Sharlin**
University of Calgary
2500 University Drive NW
Calgary, AB, Canada
1.403.210.9404
ehud@cpsc.ucalgary.ca

**Abstract.** Mouse, keyboard and graphical user interfaces are commonly used in the field of human-robot interaction (HRI) for robot control. Although these traditional user interfaces (UI) are being accepted as the standard for the majority of computational tasks, their generic nature and interaction styles may not offer ideal mapping to various robotic tasks, such as locomotion and navigation. In our research we intend to explore alternative UIs that could take advantage of human innate skills of physical object manipulation and spatial perception, and overcome some of the problems associated with traditional UIs. We suggest the use of tangible user interfaces (TUIs) for HRI applications, leveraging on existing and well-learned physical metaphors for interaction with robots, and exploring new ways to tangibly control one-to-many robot group interaction tasks. In this paper we will describe our current research efforts and findings, and outline our proposed research plans.

## 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 outputs to interact with their surroundings, and are also capable of sampling and sensing physical phenomena and translating it into digital information. As technology accelerates, advanced functionalities have been added to current robots that not only enhanced their abilities to interact with a wide range of physical objects, but also grant them the ability to communicate with humans.

In the past, researchers devoted much effort into robot development, and the problem of how to enhance human operators' situation awareness [10] when controlling robots has often been overlooked. This problem magnifies especially when a human operator needs to remotely operate one or multiple robots that have low autonomy and high intervention ratio [6]. The problem can be addressed by a set of design guidelines based on empirical studies [6, 14]. Although the guidelines are valuable for improving the operators' awareness of robots and their surroundings, they may not be well supported by the traditional user interface, that is, the mouse, keyboard and graphical user interface (GUI) paradigm which are still widely used in the field of HRI (from here on we will refer to the traditional user interface as the traditional UI).

Although the traditional UI is used abundantly in human computer interaction (HCI) tasks it may not fit well with certain HRI tasks. 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. [21] 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 letter. In terms of telerobotics, the human operators have to solely rely on the image and sensor data transmitted back by the robot to determine their next operation. Constantly switching attentions 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. Also, 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, the two-dimensional traditional UI limits people's spatial abilities when interacting with three dimensional objects. It can be difficult to control a robot that is capable of moving in three dimensions, for example an unmanned aerial vehicle (UAV) using the traditional UI. [15] In order to effectively and efficiently interact with robots, we suggest exploring an alternative set of UIs to overcome the aforementioned problems, leveraging on physical and tangible interaction metaphors and techniques.

## 2 RELATED WORK

We suggest looking for alternative solutions to the traditional UI for human robot interaction by examining tangible user interfaces (TUIs). TUIs couple digital information and function with physical objects [8] allowing a virtual entity in the digital realm to be manipulated through a physical medium. TUIs make effective use of the affordances [3] of physical objects which may allow us to fuse user input and robotic functional output together. 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 an object, 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.

Research [5, 12] have 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. [11] Thus, we can take the advantage of our innate skills at observing and learning how to interact with physical objects in interface design, which may reduce the number of new motor skills an operator needs to acquire.

When remotely navigating a robot, maintaining good spatial awareness [10] is crucial to the human operator. If a physical object can be translated into a tool for navigating a robot, then the orientation and position of the object in the physical space can be utilized to provide spatial information about a robot. Furthermore, our innate abilities allow us to interact with physical objects easily. There is no specific knowledge or memorization required for us to move, manipulate, assemble and disassemble simple physical objects pointing to the great potential of applying TUIs in HRI.

Although the notion of tangible user interface has become the buzzword in the field of Human-Computer Interaction (HCI), only very few researchers related TUIs to HRI. To the best of our knowledge, the first research project that implies the use of TUIs in HRI is done by Raffle et al. in their toy application – Topobo [9]. Topobo is a constructional toy application that allows kids to assemble static and motorized plastic components to dynamically created biomorphic forms. Not only Topobo allows creative constructions, it can also replay the motions applied by users on the motorized components to animate the user creation. Another research which we think should be considered the first attempt in the field of HRI was conducted by Quigley et al [15] who utilize a physical object for controlling a mini-unmanned aerial vehicle (UAV), using a UAV shaped physical icon for controlling the roll and pitch angle of a simulated UAV. For multi-robot control, Lapides et al. [17] have recently presented a three dimensional TUI – "The 3D Tractus" that enables a single user to monitor and control a team of independent robots in 3D spatial tasks.

## 3 FIRST ATTEMPTS

In order to explore the possibility of applying TUIs to robotic control, we have designed and conducted a user study comparing the usability of generic tangible user interfaces – based on the Nintendo Wii Remote (Wiimote) and Nunchuk [16] with a generic input device – keypad in terms of speed and accuracy in two different tasks. The study includes a high-level navigation task (Figure 1) and a low-level posture control task (Figure 2), and the study result were presented in details in [2].

One of the important advantages naturally embedded in TUIs is the physical affordance that they provide. For the navigation task that we conducted in our study, we provided two Wiimotes to the study participants for controlling a Sony AIBO robot dog [20] through an obstacle course. We have used a zoomorphic-based interaction theme: a horseback riding metaphor to explain the mechanism of controlling the AIBO using a pair of Wiimotes. The participants were asked to think of the Wiimotes as a rein on the neck of the AIBO. By pulling the left Wiimote backwards, the AIBO will rotate to the left. Reversely, pulling the right Wiimote will make the AIBO to rotate to the right. Our study demonstrated that this metaphor helped participants to quickly master the navigation task.

For the posture task, the participants were asked to command the AIBO to perform a series of postures displayed on a computer monitor (Figure 2). Both of the Wiimote & Nunchuk and keypad interface utilize an asymmetric bimanual [18] interaction style.

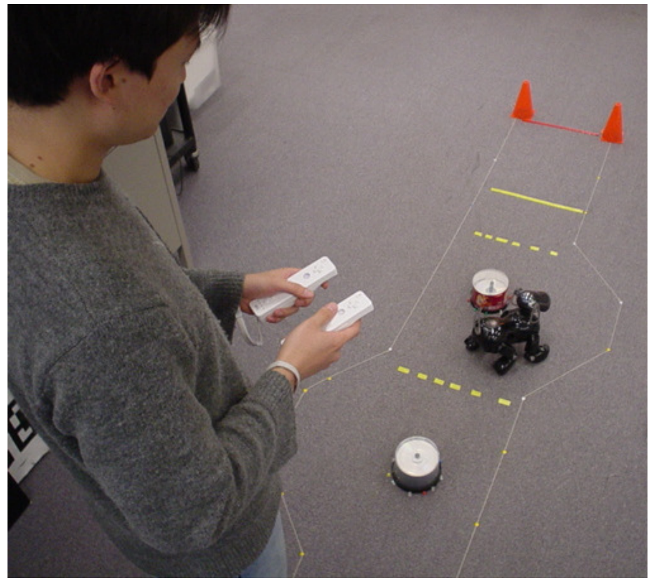

**Figure 1**. The user is navigating an AIBO robot dog through an obstacle course using two Wiimotes.


**Figure 2.** The user is controlling the AIBO to perform a posture using one Wiimote and one Nunchuk on each arm.

Due to the nature of the tasks, the Wiimote & Nunchuk gesture-to-action mappings deployed in each task differ from each other in terms of "degree of integration" and "degree of compatibility". [13] The interface mapping for the navigation task has a less than one degree of integration and a low degree of compatibility, where the interface mapping for the posture task has a close to perfect degree of integration and a high degree of compatibility.

The result of the comparative study has shown that the Wiimote and Nunchuk interface allowed the participants to finish both tasks faster, and with fewer errors than the keypad interface. Also, the majority of the participants have reported that they prefer to use the Wiimote and Wiimote & Nunchuk interface for both tasks.

This experiment suggests that using intuitive TUI-based gesture-to-robot action mapping helps the participants to reduce their cognitive load when controlling robots. This implies that operators may spend more time on high-level task planning among other tasks.

## 4 RICONS FOR ROBOTIC GROUP CONTROL

Our next step is to find a specific set of tools and interaction metaphors to design a tangible user interface for remote control of multiple robots. We intend to explore the possibilities of using small set of physical objects which resemble the shape of real robots as Ricons (robotic icons, based on Ishii & Ullmer's "Phicons" [8]) to provide a physical handle to an operator for interacting with multiple robots remotely.

## 4.1 DESIGNING RICONS

First of all, an appropriate Ricon should provide a tight spatial mapping [4] between itself and a real robot. As mentioned earlier, the shape, size and weight of a Ricon should reflect the physical properties of the robot it represents. Also, it is important and beneficial if we can utilize the physical constraints of the Ricons to prevent navigation accidents from happening. One obvious example is that each Ricon occupies a portion of the physical space. Thus, two Ricons can never "collide into" each other. This physical constraint can be immediately perceived by the operator if two robots are about to collide. Secondly, by manipulating a Ricon directly, the human operator should be able to adjust the position and orientation of a single or group of robots. For instance, when a robot or a group of robots needs specific attention, the operator can use a Ricon to give specific movement orders to one or multiple robots that are of the same type. The operator can simply move a Ricon or rotate it on a 2D surface to move or rotate a robot in the 3D space. Thirdly, the operator can use Ricons to configure different group formations of multiple robots. Multiple Ricons can be placed at different locations on a 2D region to represent the team formation of multiple robots.

To aid the human operator with sensory data and live video feedback from the robot, we want to utilize a digital tabletop for displaying such information. As Yanco et al. suggested in their research [7], to increase the operator's situation awareness in HRI interface design, we need to 1) fuse all related information onto the same display window, 2) provide spatial information in regard to the environment that the robot is within. To follow this guideline, we intend to project sensory data and live streaming video of each robot onto the digital table. In addition, to support the operator with spatial information, we can project a digital map (if available) of the remote region that the robots are working at on the table as well.

In order to closely combine the digital information with the Ricons together, we intend to put the Ricons on top of the digital table and use a vision tracking system to keep tracking of their locations on the table. By accurately locating the whereabouts of the Ricons, we can "superimpose" the Robotic status associated with each Ricon beside it. In addition, if we can access the location of each robot in the real world using vision or GPS tracking, then by scaling the digital map properly, we can use the Ricons to pin-point each robot on the map and control them in the real world by simply moving the Ricons TUI-representations on the table. This hybrid interface will not only allow I/O unification on the same surface, but also provides the ability to the operator to interact with digital and physical entities at the same time.

To simulate a robot collaboration task in a lab setting, we intend to use five to eight AIBOs as the robotic platform for performing a set of collaborative tasks. For instance, the robots will be placed in a particular formation to carry or pull a heavy object together from one place to another. (Figure 3)

**Figure 3**. A conceptual design of a simple collaborative task among AIBOs carrying an object from one location to another.

Figure 3 demonstrates one possible example of group collaboration tasks among the AIBOs. For completing tasks like this, the AIBOs have to maintain a particular group formation while moving towards their destination. If any member of the AIBO group falls behind the others, they may drop the object they carry, which in turn, fail the task.

## 4.2 SYSTEM IMPLEMENTATION

We intend to use small dog-shaped toys as Ricons TUIs for controlling the real AIBOs. By placing reflective markers on top of these toys, we will be able to use the Vicon MX system [21] to keep track of the Ricons' locations on a SMART board [19]. (Figure 4) As the users move the Ricons around on the board, the information provided by each robot will be displayed and follow along with the Ricons.

**Figure 4.** The user is holding two physical objects to interact with the virtual entities displayed on the SMART board. The board is surrounded with six Vicon motion tracking cameras for locating the reflective markers attached on top the objects to approximate their positions on the table.

In order to access the location of each AIBO in the real world, we will use another set of our lab's Vicon MX cameras to keep track of the AIBOs at a remote place (a different location from the Vicon & SMART board setup to simulate a remote robot control environment). As AIBOs move around the real world, their status and locations will be gathered and updated on the SMART board.

Since we are designing a group interface for controlling multiple robots, we are considering a layer of specific physical tools on top of the Ricons to address some of the group task aspects. In order to allow multiple robots to march in a particular formation, we intend to utilize different types of physical "ties" to accomplish this task. We define a tie as a rigid band that bounds multiple Ricons together in a pre-defined shape. For instance, reflecting on the triangle used for Pool or Billiard balls, we may build a triangle shaped tie to band multiple Ricons together in a triangle formation. By pushing the tie, we can navigate a group of Ricons to a desired location in a triangle formation easily. We may also build various ties in different shapes for accomplishing different tasks. On the other hand, by simply taking off a tie from a group of robots would break their group relationship. We hope this simple physical "binding" and "unbinding" metaphor would help users to organize multi-robot group behaviors easily.

## 5 CONCLUSION

We believe low-level robotic control tasks can benefit from the physical interaction style afforded by TUIs. The idea of using Ricons as physical handles for controlling real robots can hide tedious low-level robotic control mechanism from the end user. Moreover, the users are not required to learn new motor skills to control complex robots. By leveraging the advantage of TUIs, we can reduce the cognitive load of the human operator and allow them to spend more time on high-level task planning.

Although the human operator can directly manipulate real robots using Ricons, they can not visualize the internal state of the robots from observing the Ricons. To augment the Ricons with the information in regard to the internal status of the robots, we will use a digital table for displaying such information to aid the operator in remote control tasks. By fusing the system input and output within the perceptions of the users, we hope to reduce confusions in regard to inadequate situation awareness problem found in previous research [10].

During the development of our proposed project, we intend to explore possible physical metaphors to extend the users' ability to interact with the system based on previous knowledge. For instance, the "tie" example that we explain in Section takes advantage of people's knowledge about physical objects to easily group or separate multiple robots.

Although TUIs can provide many advantages over traditional UIs, they may be more prone to unintended usage due to their physical nature. For instance, since Ricons can be easily moved around on the table surface, users may accidentally knock them off from their supposed positions while manipulating other Ricons. Thus, we need to consider how to apply physical constraints to the system to prevent undesirable actions.

In summary, we propose to utilize both tangible user interfaces and a digital table to allow an operator to remotely navigate multiple robots. This hybrid interface will allow human operators to control individual robot behaviors and uniform group behaviors easily through the use of physical *Ricons*. No specific training will be required to operate a large robotic group with this interface. We hope our future work on the proposed system will provide new insight on human robot interface design using TUIs, especially for one to many robot navigation tasks.

## 6 FUTURE TUI DESIGN FOR HRI TASKS

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 we used in our AIBO navigation task. 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 [1] 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.

## REFERENCES


[1] BigDog, http://www.bostondynamics.com/content/sec.php?section=BigDog

[2] C. Guo, E. Sharlin. Exploring the Use of Tangible User Interfaces for Human Robot Interaction: A Comparative Study. In *Proc. CHI* 2008. Apr 5 – 10, Florence, Italy. To Appear.

[3] D.A. Norman. *The Psychology of Everyday Things.* BasicBooks. (1998)

[4] E. Sharlin, B.A. Watson, Y. Kitamura, F. Kishino, Y. Itoh. On Tangible User Interfaces, Humans and Spatiality. Personal and Ubiquitous Computing, Springer-Verlag, 2004, pp 338-346.

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

[6] H. A. Yanco, J. L. Drury. Classifying Human-Robot Interaction: An Updated Taxonomy. *IEEE* Conference on Systems, Man and Cybernetics, (2004).

[7] H. A. Yanco, J. L. Drury, J. Scholtz. *Beyond Usability Evaluation: Analysis of Human-Robot Interaction at a Major Robotics Competition*. Journal of Human-Computer Interaction, Volume 19, Numbers 1 and 2, pp. 117-149, (2004).

[8] H. Ishii, B. Ullmer. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Proc. CHI* 1997, ACM Press, 234-241, (1997).

[9] H. Raffle, A. Parkes, H. Ishii. Topobo: A Constructive Assembly System with Kinetic Memory. In *Proc. CHI 2004*, ACM Press, 647 – 654, (2004)

[10] J. L. Drury, J. Scholtz, H.A. Yanco. Awareness in human-robot interactions. In *Proc.* IEEE International Conference on Systems, Man and Cybernetics 2003. (2003).

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

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

[13] M. Beaudouin-Lafon. Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In *Proc. CHI 2000*, ACM Press, 446-453, (2000).

[14] M. Goodrich, D. Olsen. Seven principles of efficient human robot interaction. In *Proc.* IEEE International Conference on Systems, Man and Cybernetics, 2003, 3943–3948.

[15] M. Quigely, M. Goodrich, R. Beard. Semi-Autonomous Human-UAV Interfaces for Fixed-Wing Mini-UAVs. In *Proc. IROS 2004*

[16] Nintendo Wii Controllers, http://wii.nintendo.com/controller.jsp

[17] P. Lapides, E. Sharlin and M. Costa Sousa. “Three Dimensional Tangible User Interface for Controlling a Robotic Team”. Full paper in Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction 2008. HRI ’08, to appear.

[18] R. Balakrishnan, K, Hinckley. Symmetric bimanual interaction. In *Proc. CHI 2000*, ACM Press, 33-40, (2000).

[19] SMART Board ™, http://www2.smarttech.com/st/en-US/Products/SMART+Boards/default.htm

[20] Sony AIBO, http://www.sony.jp/products/Consumer/aibo/

[21] S. Faisal, P. Cairns, B. Craft. Infoviz experience enhancement through mediated interaction. In *Proc. ICMI 2005*, ACM Press (2005), 3-9.

[22] Vicon MX System, http://www.vicon.com/products/viconmx.html