Exploration of Alternative Interaction Techniques for Robotic Systems

Takeo Igarashi ▪ *The University of Tokyo*

Masahiko Inami ▪ *Keio University*

Human-robot interaction (HRI) approaches typically fall into one of two categories. One is an *agent approach*, where the user provides simple abstract instructions to the robot with voice or gesture commands, such as by saying, "Go there." In response, the robot intelligently makes one or more detailed decisions. This approach minimizes user intervention, but it often makes it difficult to offer detailed control such as to specify the path to take. The other is a *direct-operation approach*, where the user sends detailed control commands to the robot using a joystick or control pad, such as "move forward" and "turn left." Although this allows detailed control, it requires significant user attention throughout the operation.

Based on this observation, as part of our work on the Japan Science and Technology Agency's ERATO Igarashi Design Interface Project (www.jst.go.jp/erato/igarashi/en/), we have been exploring alternative approaches that fall between these two extremes, leveraging the knowledge and methodologies developed in the human-computer interaction (HCI) field. For example, the success of GUIs confirms the effectiveness of direct interaction with graphical representations using a pointing device. In turn, augmented reality (AR) techniques have validated the effectiveness of graphical overlays on top of real-world camera images, and tangible user interfaces have demonstrated the importance of direct interaction with the surrounding physical environment.

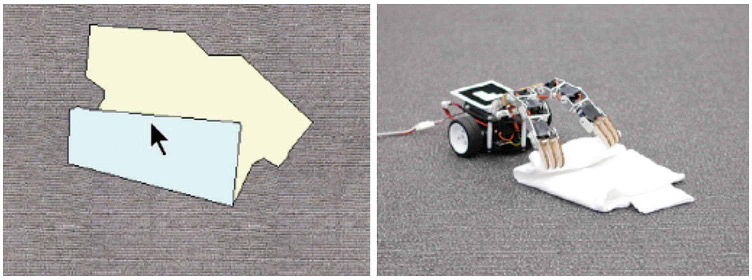
This article presents some of our explorations in this direction. Topics include GUIs for mobile robot instruction, AR methods for home appliance control, and tangible user interfaces for providing instructions to mobile robots. We also introduce sensors to enhance physical interactions with robotic systems. Our lessons learned from these experiences are determining the directions of our future research.

High-level control methods that use gestural or speech commands are overly ambiguous or excessively detailed for daily use. Human-computer interaction techniques such as GUIs, augmented reality, and tangible user interfaces could enhance physical interaction with robotic systems in the home environment.

HRI and Teleoperation

Human-robot interaction is an established field with several specialized venues such as the ACM/IEEE Human-Robot Interaction (http://humanrobotinteraction.org) and Social Robotics (www.icsoro.org) conferences. However, the basic concept underlying many of the works presented in these fields are based on the agent approach, learning lessons from human-human interaction and applying them to HRI, such as using gaze direction to express subtle communication cues. One example is Geminoid, a robot that looks and behaves like a person.[1] The principal target applications are communication robots, and few target simple robotic systems that can aid daily physical life, such as, in our case, robotic home appliances.

The related field of teleoperation introduces user interfaces to control mobile robots in remote



(a) (b)

Figure 1. Foldy, a garment folding robot. (a) GUI and (b) folding robot.

locations. One example is tele-existence,[2] where the operator controls a remote robot as if controlling his own body. When the user looks to the right, the robot looks to the right. If the user raises his hand, the robot raises its hand. These methods allow detailed control, but they require continuous, full user attention and are not appropriate as a method to provide instructions to nearby robots that assist our daily life.

In this work, we apply the methods and techniques developed in the HCI field, mainly techniques developed for real-world interaction, to interactions with robotic systems in home environments. Early works in the HCI field were designed exclusively for desktop computer systems with a mouse, keyboard, and display. As the computer has become smaller and more mobile, the focus has shifted from such desktop interactions to physical interactions leveraging advanced sensors and displays. However, most of these techniques are designed for information appliances, assisting the acquisition and control of information. They are not designed for robotic systems that physically complement our life in the real world.

Graphical User Interfaces

We initiated our exploration by implementing GUIs to teach mobile robots to perform physical tasks. Unlike speech interaction or control pads, GUIs that use a display and mouse permit the user to interact with graphical representations of the problem, facilitating efficient solutions.

Foldy: Teaching Garment Folding Graphically

Garment folding is a tedious chore in our daily life, so it would be desirable if a robot could fold garments automatically. There are several experimental systems to achieve this, but their focus is on the physical folding capability and not the user interface for instructing the robot on how to fold a specific garment. This is important because every user has a preferred manner of folding various garments. Thus, ideally, a user should be able to easily specify how to fold a garment.

Garment folding is an inherently graphical problem. Speech and gestural interactions are inappropriate for specifying a folding procedure. Control pads could allow the user to specify the folding procedure in detail, although it is tedious to continuously control a robot. Consequently, we developed a dedicated GUI to help users teach personalized garment-folding procedures.[3]

Figure 1 presents an overview of the process. The user first places the unfolded target garment on a stage and captures its image using a ceiling-mounted camera. The user then folds the garment using mouse drags, grabbing a point on the garment and dragging it to the target location to create a fold. When the user is satisfied with the result, the folding sequence is stored as a predefined procedure. The user can then instruct a folding robot to fold a garment using this procedure at a later time.

An important feature of the system is that it provides the user with visual feedback during the virtual folding process. The system continuously analyzes the validity of the current fold during direct manipulation and provides a warning to the user when a fold is invalid. For example, the lifted portion of the garment could exceed the robot's capacity. This continuous feedback facilitates rapid exploration of valid fold patterns. This would be tedious if the user were required to provide instructions using speech or a control pad.

Cooky: Cooking with Robots

In this application, a robotic system cooks a meal in a kitchen.[4] It consists of multiple small mobile robots working on a kitchen countertop and a computer-controlled heater (see Figure 2). The system pours ingredients into a pot on a heater using the mobile robots and heats the pot by controlling the heater. The user provides the system with instructions using a GUI. The interface presents a timeline representing the cooking procedure, and the user drags and drops icons representing the ingredients to specify the time to add them to the pot. The user can also specify the temperature of the heater by drawing a graph at the bottom of the display. After giving instructions, the system automatically cooks the meal while the user is working on other activities.

A key challenge is establishing an association between data in the system and the physical materials (ingredients) in the real world. To achieve this, the system uses 3D-printed, custom-made trays to locate the ingredients. The user first prepares the ingredients and places them on the special trays. A visual code is placed on top of each tray and the system recognizes it using a ceiling-mounted cam-

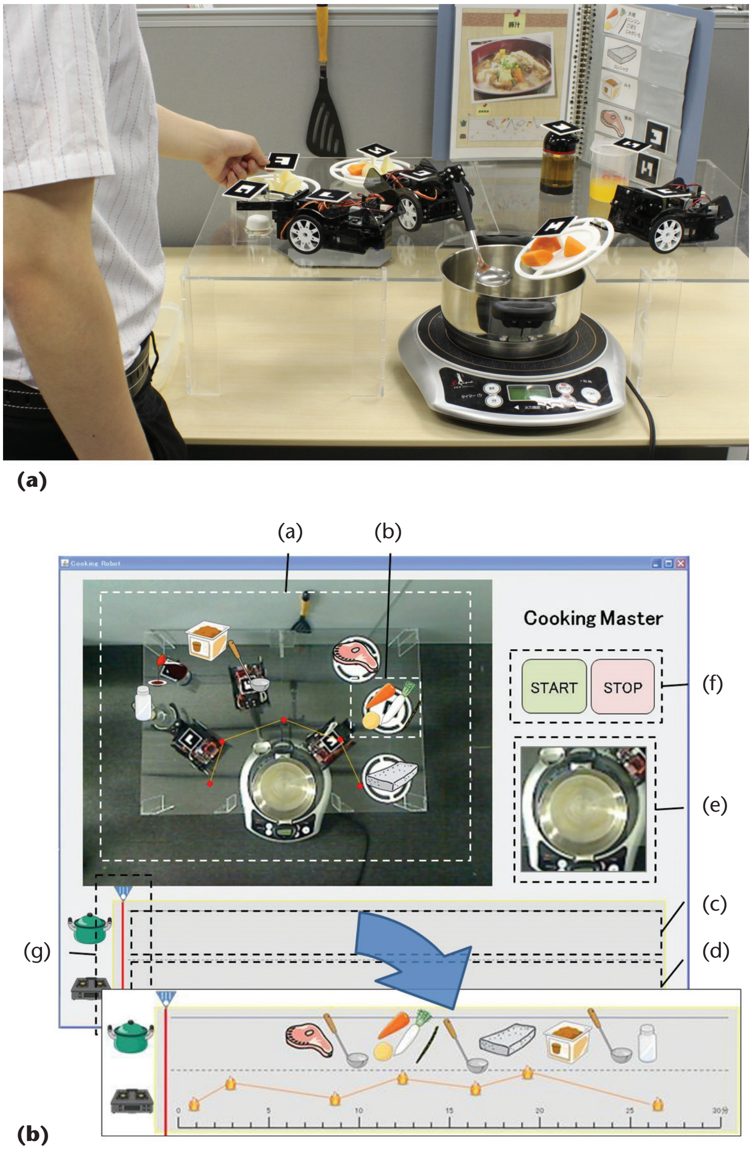


Figure 2. Cooky, a cooking robot system. (a) Robot system and (b) GUI.

era. The form factor of a tray is designed in such a manner that the mobile robot can easily manipulate it. This tray allows the system to establish an association, and the system automatically adds the ingredients to the pot at the correct time.

Augmented Reality

Augmented reality overlays virtual information on a real-world camera view to assist with a user activity in the real world. In traditional AR scenarios, the primary purpose is to provide information concerning real-world objects in the view, such as assembly instructions or geolocation information. We use AR to help the user control real-world objects remotely. This is advantageous because the user can directly interact with the target in the camera view and can see the resulting action in the same view. A classic example of applying AR to robot control is the direct manipulation of a remote robotic arm.[5] In this environment, the user manipulates the robot itself. In the proposed approach, the user manipulates the target objects in the AR environment.

CRISTAL: Tabletop Remote with Augmented Reality

The objective of this application is to help a user manage multiple home consumer electronics remote controls for devices such as TVs and digital photo frames. To achieve this goal, we capture a top-down view of a room from a ceiling-mounted camera and project the image onto an interactive tabletop surface.[6] The user then interacts with the electronic appliances in the view by directly touching them on the surface (see Figure 3). For example, the user can touch a lamp displayed in the camera view to turn on the lamp or drag a movie file and drop it on a TV screen in the camera view to play the movie on the screen. The system can also provide additional controls by displaying graphical widgets such as a slider near a target appliance. For example, the user can adjust a lamp's brightness by using a nearby slider.

We have also implemented a sketching interface for controlling a mobile cleaning robot (Roomba) using a tabletop device. When the user wants the cleaning robot to attend to a particular spot in the room, the user directs the robot by drawing a freeform stroke from the robot to the target spot to indicate a preferred user route. Lassoing using a freeform stroke is also useful for specifying target areas of variable size and shape.

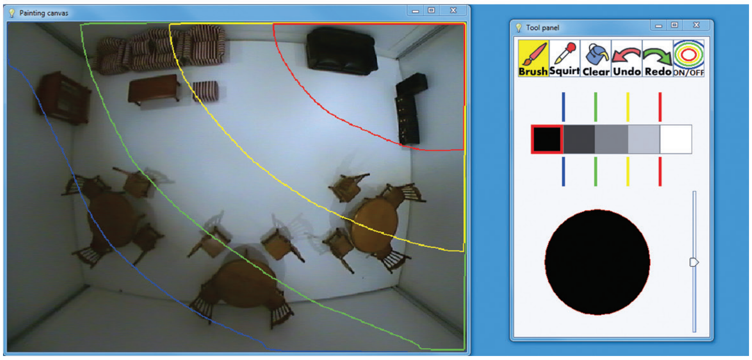
Lighty: Painting Interface for Robotic Lights

Lighty is an AR-based painting interface that enables users to design an illumination distribution for an actual room using an array of computer-controlled lights.[7] Users specify an illumination distribution for the room by painting on the image obtained by a camera mounted in the room. The painting result is overlaid on the camera image as contour lines of the target illumination intensity (see Figure 4).

The system executes an optimization process interactively to calculate the light parameters to deliver the requested illumination condition and drives the actual lights according to the optimization result. The system uses a simple hill climb for the optimization. At each iteration, it slightly changes a parameter and compares the resulting illumination condition with the requested illumination condition. The system tests multiple parameter changes and picks the one that produces a result closest to the request. We use a

Figure 3. CRISTAL, a remote control system using augmented reality. Users can control various consumer electronics remotely using the touchscreen or graphical widgets, such as a slider.

(a)



(b)

Figure 4. Lighty, an illumination control system by painting. (a) Miniature prototype and (b) painting interface.

GPU implementation of an image-based lighting simulation to estimate the resulting illumination during optimization.

We use an array of actuated lights that can change the lighting direction to generate the requested illumination condition more accurately and efficiently than static lights. We constructed a miniature-scale experimental environment (see Figure 4a) and conducted a user study to compare the proposed method with a standard direct manipulation method using widgets. The results indicated that the users preferred our method for informal light control. An interesting observation is that the proposed method was particularly useful when the user wished to make a specific location dark. Making a specific location bright is easy with standard direct control of light parameters, but making a location dark is difficult because the user must control multiple lights in a coordinated fashion. In the proposed system, making a specific location dark can easily be specified by painting with a dark brush.

Tangible User Interfaces

Tangible user interfaces employ graspable physical objects as a means for interacting with virtual information.[8] Early systems simply used physical objects such as a handle to manipulate the position and orientation of virtual objects on a tabletop environment. Later systems also used physical objects as displays or two-way I/O devices. However, their principal application continues to be interaction with virtual information, such as interaction with remote people. We use tangible user interfaces as a means to provide in-situ instructions to robotic systems that perform physical tasks in the real world.

Magic Cards: Robot Control by Paper Tags

Assume that a user wants a robot system to perform numerous household tasks during the day, such as clean a room and move the trash bin to a particular location. One possible approach is to use speech commands. However, speech is not effective for specifying a task's target locations, such as where to clean. In our proposed approach, the user places paper tags specifying a desired task at a target location in the environment.[9] For example, the user places a "vacuum here" tag at the location where she wants a vacuuming robot to clean. Similarly, the user can place a "move this object to location A" tag near a target object to move and a "location A" tag at the destination.

Once the user places the necessary paper tags in the environment, she can leave the house and the system will begin working on the tasks (see Figure 5). A visual marker is printed on the surface of each tag, and the system recognizes the tag IDs and locations using a ceiling-mounted camera. First, a tag pick-up robot collects all the tags. Then, the system executes the tasks based on the instructions left by the user. We used standard Roomba robots for the vacuum cleaning and object transport.

An important feature is that we use paper tags not only as inputs to the system but also as outputs (feedback) from the system. For example, if the system fails to complete a task for any reason (because of a low battery, for example), the system can report the failure to the user. When this occurs, the system sends a printer robot (a Roomba robot that includes a mobile printer) that leaves a paper tag with a printed error message. In this manner, the user can interact with the robot system using paper tags only, without the need for control displays or switches. This is especially advantageous for elderly users who are less comfortable touching computing devices.

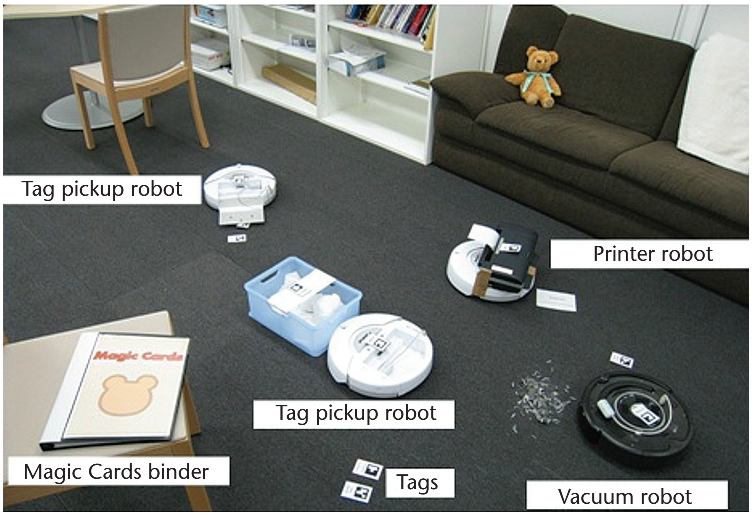


Figure 5. Magic Cards, a robot control system using paper tags. A tag pick-up robot collects all the tags left by the user, and then the robot control system executes based on the user's instructions.

Pebbles: A Tangible Tool for Robot Navigation

If it is necessary for a mobile robot to visit multiple places in a house to perform tasks, such as to deliver food to a distant room, the user should be able to specify the locations to the robot. One method is to build a map and assign a label to each location on the map, but this can be tedious and time consuming. Fully automatic methods exist where the robot automatically navigates through the environment and builds a map. However, this can also be time consuming. Another approach is to manually guide the robot to the destination using a control pad, which can be tedious.

We propose using physical, active markers as the user interface to label the environment and thus guide the mobile robot.[10] The user places landmarks, called pebbles, on the floor to indicate navigation routes to a robot (see Figure 6). Using infrared communication, the pebbles communicate with each other and automatically generate navigation routes. During deployment, the system provides feedback to the user with LEDs and speech, allowing the user to confirm that the devices are appropriately placed for the construction of the desired navigation network. Moreover, because there is a device at each destination, the proposed method can name locations by associating a device ID with a particular name.

Compared with autonomous mapping methods, this method is advantageous because it is much faster for a person to place landmarks in the environment rather than have a robot navigate through the environment. It is also beneficial that the user can encode semantic knowledge about the environment by placing the devices. For example, if a user does not want the robot to use a specific route, the user can prevent this by not placing devices in the route. Such user intention is difficult to represent in a fully automatic approach. Although it is certainly possible to provide labels using a GUI after obtaining an environment geometry, it is more efficient to select and move physical devices in the environment than activate a computer and operate a graphical map management program.

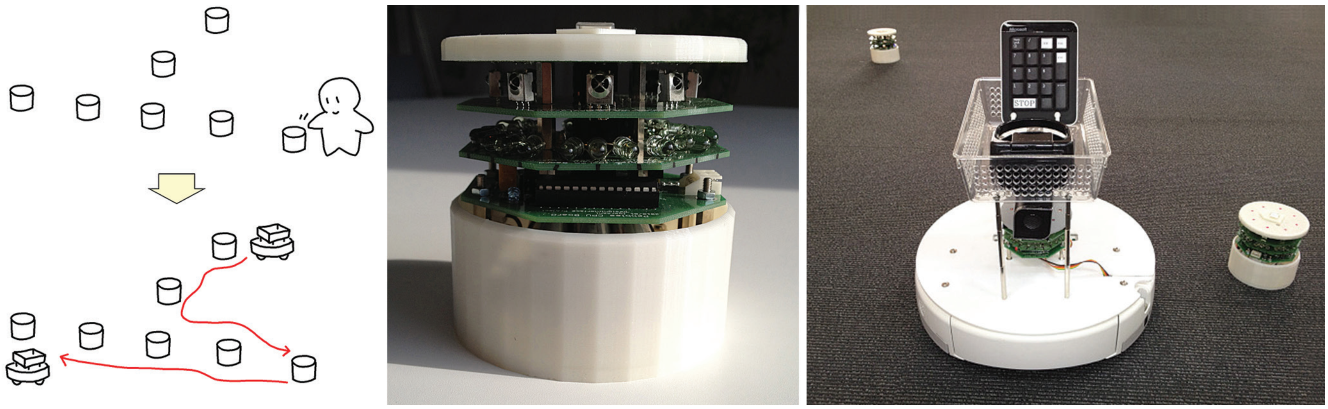
Novel I/O devices

An important trend observed in the HCI research field is the additional focus on novel hardware devices. Traditional techniques have relied solely on standard input devices (such as a mouse and keyboard) and output devices (such as LCD displays). However, we now see an abundant variety of I/O devices in use, such as pressure sensors, light sensors, haptic displays, and scent displays. This trend is motivated by the need for interaction with the real world and enabling technologies such as Arduino (www.arduino.cc) that allow the rapid development of such novel hardware devices.

To realize novel methods for interacting with the real world, we developed several dedicated I/O devices. Our primary objective is to bring nonintrusive computational and robotic elements into people's lives. We have developed various techniques to introduce softness into computing devices as an approach to achieve this goal.

FuwaFuwa: Sensing Cushion Deformations

FuwaFuwa is a sensing device for detecting the deformation of soft objects such as cushions and plush toys.[11] The popular method for detecting such deformation is to attach pressure or deformation sensors onto the object's surface. This, however, negatively affects the softness of the object. Therefore, we developed a sensor module



(a) (b) (c)

Figure 6. Pebbles, a tangible tool for robot navigation. (a) Concept and (b) implementation.

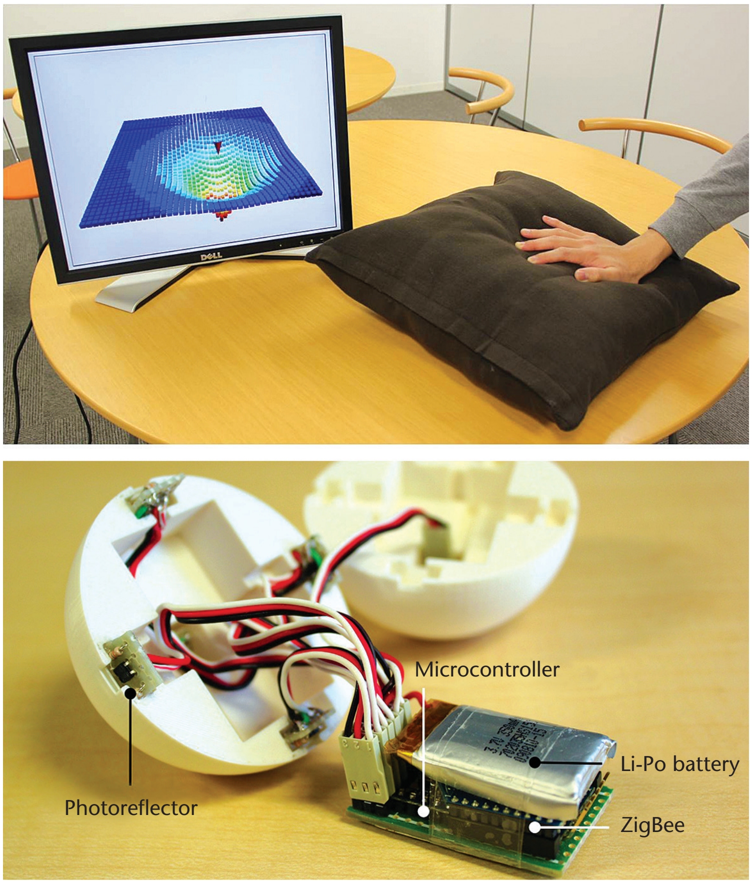


Figure 7. FuwaFuwa, a sensing device for cushion deformation. Detecting deformation from inside preserves the object's softness.

that can detect deformation from inside. We use a photo reflector that is a combination infrared light emitter and photo sensor (see Figure 7). The light emitter radiates light that is reflected by the stuffed material inside the soft object. The light intensity is measured by the photo sensor. As the user pushes the soft object, the density of the stuffed material increases. This increases the intensity of the reflected light observed by the photo sensor.

We have developed several sample applications using this device, including a game controller, media controller, and musical instrument. We also developed a soft robot that moves in response to the user's squeezing action. An important aspect of this work is the emphasis on softness in computing objects. Typical hardware devices are literally hard, consisting of metals or plastics. However, in daily home life, many soft objects come into contact with the body, such as cushions and pillows. Technologies that can convert such soft objects into interactive devices are essential to bringing technology closer to our daily lives.

PINOKY: Animating a Plush Toy

The FuwaFuwa device is a purely input device. PINOKY, a device attached to a plush toy, can detect deformation and move the toy's limbs.[12] PINOKY is a wireless ring-like device that can be externally attached to any plush toy as an accessory (see Figure 8). The device consists of a microcontroller, motors, photo reflectors, a wireless module, and a battery. The motors generate forces that move the plush toy surface sideways, forcing the plush toy limbs to bend. The photo reflectors sense the deformation of the plush toy. This input and output combination makes it possible to record motion caused by manual deformation and then reproduce the motion using the motors. We have developed various applications including remote tangible communication, storytelling, and physical toys as external displays for video programs.

An important feature of the device is that it can be attached to an existing plush toy externally, rather than embedded in the toy. Embedding a device in an existing toy is often not acceptable because it is necessary to cut the toy. Instead, the external attachment mechanism lets a user convert any plush toy into an interactive robot in a nonintrusive manner, without having to alter the toy. After playing with the toy, the user can remove the device from the toy quickly. We believe that devices that can be externally attached to an existing static object to convert it into an interactive entity are important to introducing robotic technologies into our daily lives.

Graffiti Fur: A Carpet as a Display

The last application is an output device that can convert regular carpets (which we can consider as "fur" with fibers) into a computational display.[13] This utilizes the phenomenon whereby the shading properties of fur change as the fibers are raised or flattened by a finger. It is possible to erase drawings by first flattening the fibers by sweeping the surface by hand in the fiber's growth direction and then draw lines by raising the fibers by moving a finger in the opposite direction.

We have developed three different devices to draw patterns on a "fur display" utilizing this phenomenon: a roller device, pen device, and pressure projection device (see Figure 9). The roller device has an array of rods underneath. These rods move up and down independently as the user moves the device on the fur. Lowered rods selectively raise the fibers, leaving a pattern on the surface. The pen device is used for freehand drawing by a user. A small continuously rotating rubber wheel is attached to the pen tip, and it raises the fibers when in contact. The pen device is also equipped with a gyro sensor and continuously adjusts the wheel's orientation such that it can raise the fur regardless of the holding posture. The pressure projection device uses focused ultrasound to remotely raise or flatten the fur.

This technology can convert ordinary objects in our environment into rewritable displays without requiring or installing nonreversible modifications. More importantly, this display technology does not require energy to maintain the imagery appearing on the display. This energy-saving feature is important because it promotes the reduction of energy consumption at home.

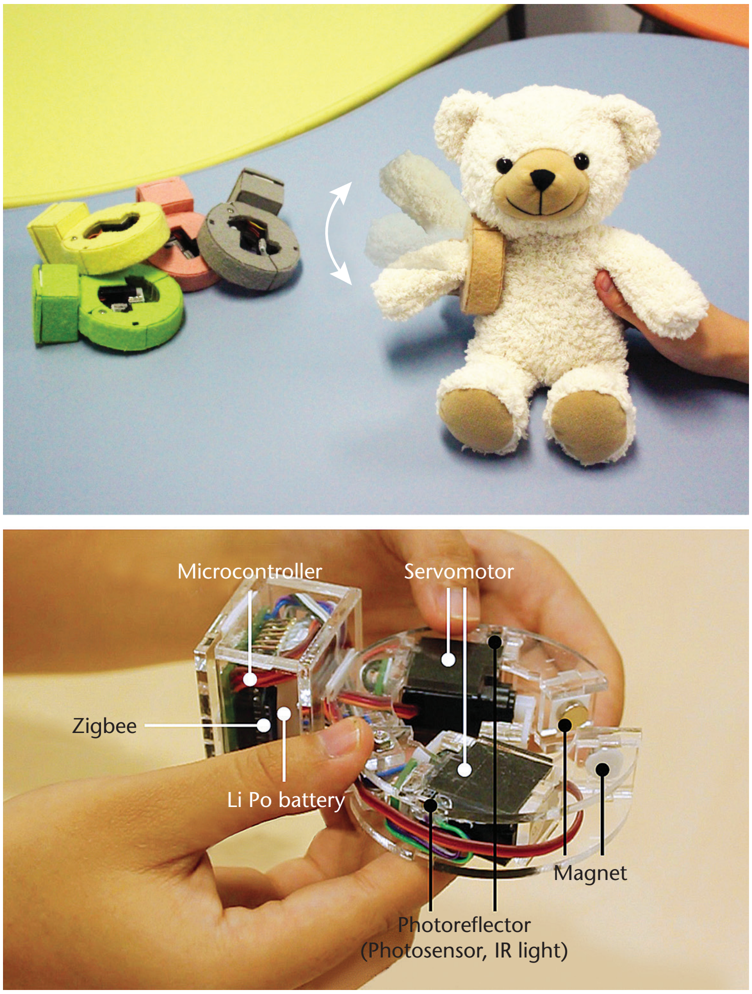


Figure 8. PINOKY, a device for animating a plush toy. Instead of an embedded device, PINOKY is an external attachment that can convert any plush toy into an interactive robot by recording motion caused by manual deformation and then reproducing the motion using motors.

Discussion

We launched this project to identify interaction methods for robotic systems that address the limitations of traditional intelligent-agent approaches and full low-level controls. Having completed several development projects and experiments including those introduced in this article, we now believe we should strive for transparent or implicit user interfaces when designing user interfaces for robotic systems in home environments. A transparent user interface allows users to interact with the real world directly, without being aware of the computational system between them and the real world. This is similar to what GUIs have attempted to offer. However, traditional GUIs are designed to manipulate information in the virtual world, and transparent interfaces aim to manipulate the real world. Representative works embodying transparent interfaces are those based on AR. Systems that use GUIs and tangible interfaces can also be seen as methods of allowing users to interact with the real world while concealing the low-level control interface required for robotic systems.

To enable such transparent user interfaces, it is critical to establish correspondences between objects in the real world and their virtual representation in the interface. We employed several methods such as fixed screen coordinates in the view from fixed cameras, visual markers, and electronic signals. More efficient and nonintrusive methods must be developed in the future. Computer vision techniques are rapidly advancing, and it is now possible to obtain a full 3D geometry of

Figure 9. Graffiti Fur, devices for drawing on a carpet. (a) This technology can convert ordinary objects into rewritable computational displays using (b) a roller, pen, or pressure projection devices.

the surrounding environment in real time. However, obtaining a 3D geometry is not sufficient for building robotic systems that execute tasks for the user. The challenge is assigning meaning to this geometry. Sophisticated interaction techniques are necessary because the meaning of environment is different for every person, and ultimately only the individual user can define this meaning.

Evaluating interaction methods developed in this project was difficult. These interaction methods are designed for futuristic (nonexistent) robotic systems and therefore direct comparison is not possible because there is no baseline method. In addition, our work is more like presenting new applications than improving interaction for existing applications. So, most evaluation results are qualitative; we asked test users to try the prototype systems and collected feedback and suggestions for further improvement. Most of them provided us with positive feedback, saying that they want such applications when they are available. For example, in the Magic Card test, users appreciated the ability to interact with robotic systems with paper cards, saying that it might be especially appreciated by people with technophobia who dislike interaction with devices (buttons or screens). At the same time, some expressed concern that the method might be problematic for families with little children. As for PINOKY, most test users reported that the device was enjoyable and easy to use and that they did not feel the need for extensive system training. The device was difficult for two- and three-year-old children, but elementary and junior high school students were able to utilize it without a problem.

Another important lesson from this effort is that it continues to be difficult to build a system that executes tasks involving the manipulation of physical objects, such as carrying an object from one place to another. To grab and transport an object, you must include a powerful arm and mobile base. This results in a device that is overly expensive and/or too bulky for a home application. Consequently, we gradually migrated our focus from general-purpose mobile robots to the enhancement of existing home appliances, such as lights and electric fans. These intelligent appliances are available today and require effective user interfaces. As less expensive and more efficient motors and sensors become available, we foresee that more objects in a house will become intelligent, and the need to develop sophisticated interaction techniques and I/O devices such as those we introduced here will increase.

We believe that our work complements existing work on social robotics. Social robotics take an agent approach, where the user interacts with a robot using communication modalities found in human-to-human communication such as speech, facial expressions, and body gestures. Social robots must present themselves to users to provide them with social support, communicating information to the users and guiding them. Our approach, on the other hand, regards a robot as more like a tool to assist physical activities. To that end, we try to hide the robotic systems from users and allow them to transparently interact with the real world. The robotic systems in the future will probably have both aspects, combining the agent and tool approaches to provide the best interaction method for various tasks.

As a result of this work, we discovered that transparency is key for the development of intuitive user interfaces for systems that function effectively in the real world. The user must be able to experience a feeling of transparently manipulating the real world. The robotic system provides the mechanism to allow the manipulation to occur. Consequently, it is necessary to establish correspondences between real-world entities and information in the robotic system. These correspondences are dependent on the individual user and environment. Therefore, fully automatic methods are not obtainable, and sophisticated interaction methods are required. Finally, we discovered that the development of general-purpose mobile robots that execute physical tasks remains a difficult task. However, presently, intelligent robotic home appliances are being made available, and there is a significant need for advanced interaction methods for such appliances.

References

1. S. Nishio, H. Ishiguro, and N. Hagita, "Geminoid: Teleoperated Android of an Existing Person," *Humanoid Robots, New Developments*, A. Carlos de Pina Filho, ed., I-Tech, 2007.
2. S. Tachi et al., "Tele-existence (I): Design and Evaluation of a Visual Display with Sensation of Presence," *Theory and Practice of Robots and Manipulators*, pp. 245–254.
3. Y. Sugiura et al., "Graphical Instruction for a Garment Folding Robot," *Proc. ACM Siggraph 2009 Emerging Technologies* (*Siggraph*), 2009, article no. 12.
4. Y. Sugiura et al., "Cooking with Robots: Designing a Household System Working in Open Environments," *Proc. 28th Int'l Conf. Human Factors in Computing Systems* (*CHI*), 2010, pp. 2427–2430.
5. D. Drascic et al., "ARGOS: A Display System for Augmenting Reality," *Proc. INTERACT 93 and CHI 93 Conf. Human Factors in Computing Systems*, 1993, p. 521.
6. T. Seifried et al., "CRISTAL: Design and Implementation of a Remote Control System Based on a Multi-touch Display," *Proc. ACM Int'l Conf. Interactive Tabletops and Surfaces*, 2009, pp. 37–44.
7. S. Noh et al., "Design and Enhancement of Painting Interface for Room Lights," *The Visual Computer*, vol. 30, no. 5, 2013, pp. 467–478.
8. H. Ishii, "Tangible User Interfaces," *Human-Computer Interaction: Design Issues, Solutions, and Applications*, A. Sears and J.A. Jacko, eds., CRC Press, 2007.
9. S. Zhao et al., "Magic Cards: A Paper Tag Interface for Implicit Robot Control," *Proc. ACM Conf. Human Factors in Computing Systems* (*CHI*), 2009, pp. 173–182.
10. K. Ishii et al., "Pebbles: User-Configurable Device Network for Robot Navigation," *Proc. 14th IFIP TC13 Conf. Human-Computer Interaction* (*INTERACT*), 2013, pp. 420–436.
11. Y. Sugiura et al., "FuwaFuwa: Detecting Shape Deformation of Soft Objects Using Directional Photoreflectivity Measurement," *Proc. 24th Ann. ACM Symp. User Interface Software and Technology* (*UIST*), 2011, pp. 509–516.
12. Y. Sugiura et al., "PINOKY: A Ring That Animates Your Plush Toys," *Proc. 30th Int'l Conf. Human Factors in Computing Systems* (*CHI*), 2012, pp. 725–734.
13. Y. Sugiura et al., "Graffiti Fur: Turning Your Carpet into a Computer Display," *Proc. 27th Ann. ACM Symp. User Interface Software and Technology* (*UIST*), 2014, pp. 149–156.

Takeo Igarashi *is a professor in the Graduate School of Information Science and Technology at the University of Tokyo. His research interests include user interfaces and computer graphics. Igarashi has a PhD in information engineering from The University of Tokyo. Contact him at takeo@acm.org.*

Masahiko Inami *is a professor in the Graduate School of Media Design at Keio University. His research interests include physical media and entertainment technologies. Inami has a Ph.D in engineering from The University of Tokyo. Contact him at inami@kmd.keio.ac.jp.*