

# An Augmented Reality System for Teaching Sequential Tasks to a Household Robot

Richard Fung, Sunao Hashimoto, Masahiko Inami, Takeo Igarashi

**Abstract**—We present a method of instructing a sequential task to a household robot using a hand-held augmented reality device. The user decomposes a high-level goal such as “prepare a drink” into steps such as delivering a mug under a kettle and pouring hot water into the mug. The user takes a photograph of each step using the device and annotates it with necessary information via touch operation. The resulting sequence of annotated photographs serves as a reference for review and reuse at a later time. We created a working prototype system with various types of robots and appliances.

## I. INTRODUCTION

ROBOTICS technology has made a significant progress in recent years and now we can see domestic robots performing various household tasks in laboratories such as picking up garments from a washing machine, picking and folding towels, washing dishes in a kitchen, and cooking a meal. However, most of these operations are preprogrammed by the developers and are difficult to customize by end users. Since each home environment is different and each individual has his/her own way of doing things, we believe that it is crucial to allow the user to teach these robots a specific way of doing these tasks.

Example scenarios we have in mind are household chores that robots could perform automatically [1]. A laundry robot picks garments from a basket, delivers it to a washing machine, runs the washing machine, picks up the cleaned garments, and stores them in a closet. A cleaning robot moves furniture out of the way, opens the windows, dusts the shelves, vacuums the room, closes the windows, and replaces the furniture. A waiter robot opens a table by cleaning it, arranging place mats and utensils, and delivering dishes. These scenarios show tasks with sequential structures and interaction among robots, objects, and appliances.

To achieve this goal, several robot problems must be addressed. Starting from the low level, a robot needs to know how to manipulate objects and appliances. A robot also needs to know how to orient itself and move around in a domestic environment. Lastly, a robot needs to know how to complete a high-level task such as “clean my room.” The scope of this research addresses the latter problem,

instructing a robot how to perform a high-level sequential task.

We believe that a human should teach the steps of a sequential task to a robot. A robot is not likely to have the cognitive abilities of a human in the foreseeable future, but humans can decompose a complex task in such a way that a robot understands.

In this paper, we present a method of teaching a sequential task to a robot using photographs. We introduce a handheld augmented reality device to create photographic instructions. The device is similar to a digital camera with touch operation for annotating photographs. Importantly, the augmented reality device does not require the robot’s presence to create instructions.

There are several benefits in using photographic instructions. First, the teaching is situated in the actual environment. Second, the user demonstrates where to place objects. Third, the device ensures that photographs are properly composed. A properly composed photograph has a visible subject, object, and verb.

When using the device, icons are overlaid on physical objects that the robot can manipulate. These icons provide feedback about the robot’s ability to recognize objects. After an object’s icon is selected, the device lists available actions that a robot can perform on the object. For instance, at a hot water kettle, icons are shown for both a mug and hot water kettle. The user selects the kettle icon to tell the robot to heat up water, and the user selects the mug icon to tell the robot to reposition the mug. Several of these photographs are taken in sequence to form the steps of a high-level task, which is then reviewed by the user and executed by the robot.

We demonstrate this concept in a prototype system with a basic robot that can move items to a target location and operate home appliances. We tested the augmented reality device in two scenarios, preparing a drink and organizing a room. To prepare a drink, a table top robot delivers a cup to hot water kettle and other appliances to make coffee. To organize a room, a robot pushes objects on the floor.

In conducting this research, our paper makes two contributions. First, we present an augmented reality interface using a handheld device, which allows for situated teaching. That is, the user can be in the actual environment where the task occurs, making it easy to identify environmental constraints. Second, a sequential task is represented as annotated photographs. Annotated photographs are semantically meaningful to both the user and robot. The user recognizes the physical environment

Manuscript received March 17, 2011. This research is supported by the JST ERATO Design Interface Project, Japan.

R. Fung is with the University of Calgary (fungr@ucalgary.ca). S. Hashimoto is with JST ERATO Igarashi Design Interface Project, Japan (hashimoto@designinterface.jp). M. Inami is with Keio University, Japan (inami@inami.info). T. Igarashi is with the University of Tokyo, Japan (takeo@acm.org).

from the photographs, and the robot interprets the annotations within the photograph explicitly (e.g., a photographed home appliance is annotated with a “turn on” action).

## II. RELATED WORK

We discuss several strategies to specify high-level tasks to a robot such as natural language commands, demonstration, visual programming languages, virtual simulations, live video, and photographs.

The goal of this research is to enable a human to specify a high-level task. Commanding a robot with natural language commands is the ideal sci-fi solution (consider Jetsons’ Rosie or Star Wars’ 3CPO humanoid robot). Realistically, we cannot assume that robots will have the same cognition as humans and learn tasks in the way that humans do. Rather, a restricted natural language vocabulary may be used for teaching [2][3]. A natural language interface, however, requires a user to *recall* the restricted vocabulary. The augmented reality device, on the other hand, allows a user to *recognize* the available vocabulary visually.

Prior work sought how to teach robots by demonstration [4]-[6], for example, to manipulate objects by recording a person’s arm movement [7][8]. Teaching an object manipulation task is accomplished effectively by demonstration, but it becomes difficult to review the robot’s model of the learnt task.

A robot can be instructed using a visual programming language [9][10], which gives a visible representation and a recognizable set of operations. For example, the Lego Mindstorms kit and Microsoft Robotics Studio use blocks and connections to describe a robot’s sequence of actions. A storyboard of animated cartoons [11], timeline [12], and key frames [13] can be used similarly. Although visual languages provide stepwise instructions, the user programs the task away from the actual environment. The augmented reality device is a forcing function to use the actual environment for instruction, which has the benefit of recognizing environmental constraints.

Virtual simulations can approximate the robot’s operating environment. A simulation can instruct a robot to manipulate objects [14] and fold a cloth [15]. The Alice programming environment [16] demonstrates a rich set of tasks that can be accomplished in a simulation. These systems, however, require significant effort to update given that domestic environments change frequently. Photographs can be easily retaken to account for environmental changes.

Live video can be used to instruct a task to a robot. A user can manipulate a robot interactively in video by operating controls seen on the robot [17]. A user can specify where and what a robot should do in the environment using video from a ceiling-mounted camera [18][19]. Objects that are occluded from the camera’s view, however, cannot be specified using these approaches. Using an augmented reality device, which prompts the user to rearrange objects,

allows a user to reveal occluded locations.

Our work chooses static photographs over live video because of irrelevant nuances in video. Nuances introduced during a live video demonstration (e.g., hand tremor) are irrelevant to the task and shouldn’t be saved. Consider an analogy of creating animations: only key frames are necessary.

The most similar work uses a photograph to restore the layout of multiple objects on a table [20]. However, they do not consider sequential tasks that involve multiple objects and home appliances.

Our approach of building sequential tasks from steps has been considered in artificial intelligence research. For instance, a hierarchy of primitive robot operations and sequences can be used to construct robot behaviors [21]. Whereas prior work is primarily concerned with the learning algorithm, we explore the human-robot interface and propose an augmented reality interface.

## III. TEACHING WITH PHOTOGRAPHS

We first present an example scenario that uses photographs to specify a sequential task. We then describe an exploratory study of using photographs to teach another person.

### A. Example Scenarios

An ideal robot should know how to perform household chores such as preparing a cup of coffee and organizing a living room. A preprogrammed out-of-the-box solution does not work because every household has its own idiosyncrasies. Photographs allow a user to program tasks that are tailored to his/her home environment.

Let us consider two examples of teaching a robot using photographs. Suppose that a user wants a robot to prepare a cup of coffee every morning. The user takes out a handheld device (e.g., a smart phone) to teach the steps. The user first takes photographs of a mug at the different appliances such as a powder dispenser (to dispense coffee), hot water kettle, and dining table. The user annotates each of these steps to ensure that the robot understands the instructions.

In another scenario, the user wants to teach a robot how to organize a living room. The user takes photographs of objects on the floor, such as a magazine box and trash bin, that should be cleaned. Then, the user takes another set of photographs to show where they should be placed.

### B. Exploratory Observations

These example scenarios show a task involving multiple steps. We conducted a mini-study with four participants using a regular digital camera to see how people use photographs to explain a sequential task, identify objects, and instruct the operation of appliances. We asked participants to teach another person how to prepare coffee using only the camera. The environment had a coffee powder dispenser, hot water kettle, and several mugs.

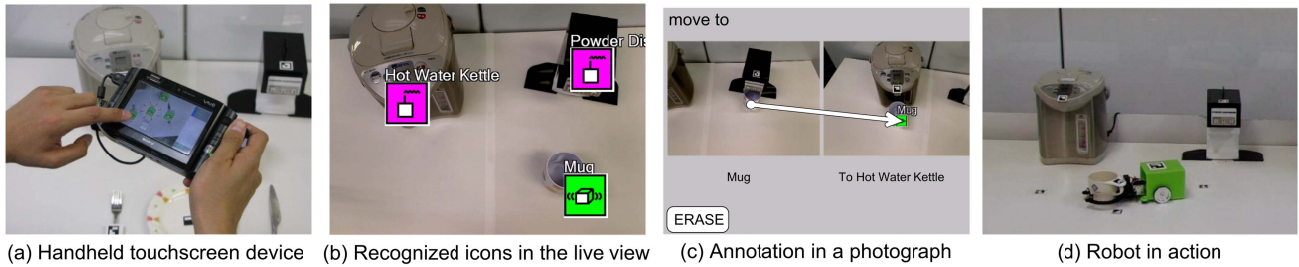


Fig. 1. A household robot is taught to deliver a mug to various home appliances and locations.

The four participants were from our research lab with a university-level education in engineering, but they had no prior knowledge of this project. The sessions were conducted in our lab’s kitchenette and were video recorded for analysis.

Participants took two strategies to identify target objects and an appliance’s buttons. One strategy was to photograph the subject of interest without context (in other words, zoom in). In another strategy, participants photographed their hand pointing at an object or button.

Three participants showed the movement of objects using two or more photographs, one photograph to show only the object and another photograph to show the object at an appliance. Two participants also photographed the object in its original location (e.g., on a shelf) before showing a close up photo of the mug.

No participant was able to use photographs to specify a time to wait. One participant wanted to add a verbal annotation. Another participant took three photographs to represent three seconds, but he wanted to write down the duration and include it in the photograph. We had expected participants to photograph a nearby clock.

We learned from this study to allow annotations on photographs using a tapping gesture. We also modeled the from-and-to photographs in our design. Some concepts such as wait were too abstract for photographs and we decided to present them in a graphical user interface.

#### IV. AUGMENTED REALITY INTERFACE

Photographs can be as precise or ambiguous as the photographer chooses. In order to constrain how photographs are taken, we provide an augmented reality device to scaffold the creation of interpretable photographs. By adding annotations to a photograph during construction, both the user and robot have the same interpretation of the photograph.

##### A. Teaching a Task

We created a handheld augmented reality device with a video-see-through mode, which we implemented using a Sony Vaio ultra mobile PC (Figure 1a). The device displays icons and captions on objects in the see-through display as shown in Figure 1b. The order in which photographs are taken is the order for completing a sequential task.

The system recognizes two categories of physical objects, noun objects and verb objects. Nouns are *moveable items*,

objects that the robot interacts with directly. Verbs are *home appliances*, a type of object that affects the state of another object.

Each photograph is taken as follows. The user creates an instruction by pointing to an object. A tapping gesture is used to annotate objects. The user can also specify actions directly to the robot without a photograph; for example, the Wait button instructs the robot to wait after the last photographed step. Next, we describe how to create instructions for moveable items and home appliances.

##### 1) Delivering Movable Objects

The handheld device can instruct a robot to move objects such as dishes and trash baskets. We implemented and tested two types of move instructions. The “move to” instruction requires two photographs, one for the identifying the object and another for the destination. We describe the interaction for delivering a mug to a table as follows (Figure 2):

1. The user frames the mug’s icon in the device’s live view (Figure 2a). The location such as a table or home appliance should also be visible in the live view.
2. The user taps the moveable item icon.
3. A photograph is taken and a list of actions appears (Figure 2b). The user chooses the “move to” action by tapping it on the display. Multiple objects can be chosen from the photograph before choosing the “move to” action.
4. The user physically demonstrates where the mug should be placed. In Figure 2c, the mug is placed at another table.
5. The user frames the live view such that the mug and a location icon are both seen. If the object is not seen in the photograph, the user is prompted to move the object to its destination and include it in the live view.
6. The user taps the live view to complete the instruction.

These steps ensure that the instruction is syntactically valid for the robot. Photographs without annotations can be interpreted in many ways, for example, a photograph of a mug and kettle can be interpreted as filling a mug or moving the mug from the kettle. The annotations ensure that there is a common interpretation between the human and robot.

A photograph requires sufficient context such that a location can be identified. Locations are presented as location icons in the live view. The device prompts the user to include more photographic context if it cannot be determined from the photograph. If multiple location icons appear on the display, the user is prompted to recompose the live view with a single location icon.

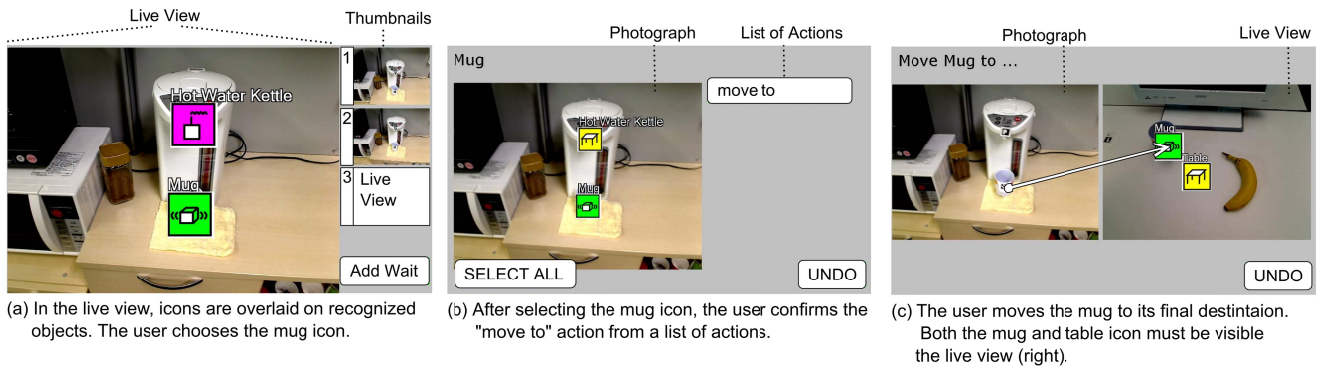


Fig. 2. Creating an instruction to move a mug from a hot water kettle to a table.

Semantic checks are performed on the instruction. If a moveable item is placed at an appliance, the device checks to ensure that an object can be placed there. If the moveable item should not be moved to an appliance, a “not allowed” icon is shown. For instance, a plate cannot be used at a hot water kettle.

### 2) Saving Object Layout

The augmented reality device provides a “move here” instruction, similar to the “move to” instruction. This instruction is provided to photograph an environment’s ideal state, which replicates prior research in saving table layout [20]. Only one photograph is taken with both object and location icons. All relevant object icons must be selected one at a time.

### 3) Operating a Home Appliance

Our system supports the operation of home appliances. The operation of home appliances is similar to controlling information appliances (e.g., lamps, radios) directly [22]. This work extends prior work to instruct a robot in operating a home appliance, which in turn modifies an object. For example, a hot water kettle pours water into a container. An instruction is created as follows:

1. The user places a target object at a home appliance.
2. The user selects the icon for a home appliance in the device’s live view (Figure 3). If the appliance requires a target object and multiple objects are seen, the closest object to the appliance is selected by default.
3. The user selects an action from a list of appliance’s actions (Figure 3). For example, a hot water kettle has the “heat”, “fill”, and “turn on” actions. If there is only one action, this step is skipped.
4. If the action has parameters, the user is presented with a sub-list. For example, the kettle’s “heat” action has a sub-list to set its temperature.

Semantic checks are performed when operating appliances. In Step 2, a valid target object must be seen in order to select an action. For instance, a kettle should pour water into a cup, but it should not pour water onto a plate; thus, the photograph must contain both the hot water kettle and cup (presumably under the kettle’s spout) before the *fill* action is made available.

### 4) Instructing the Robot to Wait

The user can tell the robot to wait between photographed

instructions using the “Add Wait” button, which is seen in the bottom right corner of Figure 2a. When the user presses the button, the user specifies the duration to wait in a graphical user interface (Figure 3).

### 5) Instructing the Robot in Different Locations

The user specifies the spatial context of a move or operate instruction by including a location icon within the live view. Locations are used by the “move here” and “move to” instructions to identify a destination. The system records the target object’s position relative to the nearest location icon, which is used when executing the task. Home appliances are also destination locations for “move here” and “move to”.

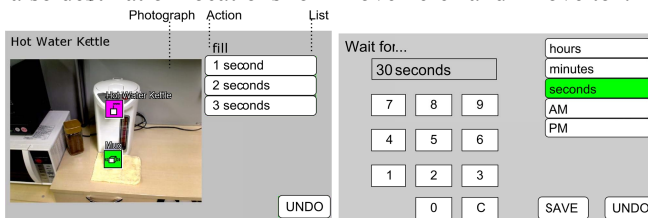


Fig 3. Instructing the robot to operate a hot water kettle (left) and wait for 30 seconds (right).

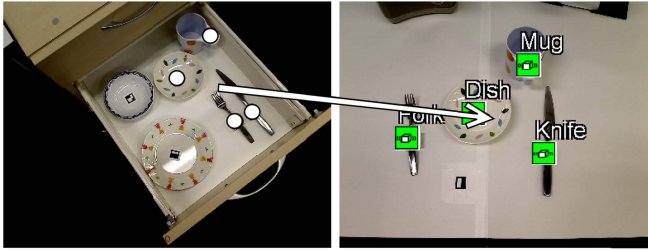
## B. Reviewing and Executing a Task

After creating sequential steps, the resulting task is displayed as a sequence of annotated photographs. Each photograph is displayed with icons, captions, and arrow lines. A caption for the whole photograph is placed in the upper-left corner to indicate an action in the photograph, for example, to “turn on” a faucet (Figure 4b). When multiple objects are being moved, a circle identifies each object in a “before” position (Figure 4a). Arrow lines indicate the movement of objects.

The annotations provide a visible representation of the robot’s internal model of the task to the user. The user recognizes that the robot will only interact with the highlighted objects in the photograph (for example, some dishes in Figure 4a are ignored).

## V. PROTOTYPE SYSTEM

We implemented a prototype system with two household robots to show the feasibility of the augmented reality device. The prototype is deployed in a structured home environment with robots and home appliances. We chose a



(a) A move operation is shown with an arrow line and icons



(b) A home appliance is shown with a caption and icons

Fig. 4. Annotated photographs.

structured environment because our primary focus is in human-robot interaction, not the robot’s capabilities.

### A. Structured Home Environment

A structured home environment has physical objects, appliances, and locations predefined in a database (Table 1). The database contains semantic information about objects such as their associated types: a moveable item, home appliance, or location. It also contains information about the available actions that can be performed on an object or at an appliance (Table 2).

Physical objects, appliances, and locations are affixed with fiducial markers for vision recognition. These fiducial markers are recognized by the augmented reality device and overhead cameras. Overhead cameras provide a 2D coordinate space for objects, appliances, and robots, which is an approach employed for robot localization [23][18].

Affixing fiducials on all objects is not the ideal solution. However, we believe that our assumptions of recognizing objects and appliances is reasonable given that there have been significant progress in pure vision recognition [24] and wireless tracking technology [25].

### B. Robots

We programmed two household robots in our home

TABLE 1. THE HOME ENVIRONMENT FOR TWO SCENARIOS.

	Scenarios	
	Prepare a Drink	Organize a Room
<b>Home Appliances</b>	Powder dispenser Pill dispenser Hot water kettle Faucet	
<b>Moveable items</b>	Mugs and cups Teaspoon Plate	Magazine bin Trash can
<b>Locations</b>	Kitchen counter Dining table	Anywhere on the floor

environment. The first robot is a custom-made robot (14x11x8 cm, Figure 1d) with a claw to grab objects up to 16 cm wide on a flat surface such as a table. We hard coded a single grabbing angle, but it is possible to compute the grabbing angle with on-board sensors. The second robot is a modified iRobot floor robot, which is used to push objects around. Given the ability of both robots, all moveable items are registered to either “move to” or “move here” actions.

TABLE 2. AVAILABLE ACTIONS AND PARAMETERS

Subject	Possible Actions
<b>All moveable items</b>	Move to Move here <sup>1</sup>
<b>Power dispenser</b>	Dispense (1, 2, or 3 seconds)
<b>Pill dispenser</b>	Small pill, medium pill, large pill
<b>Hot water kettle</b>	Fill (1, 2, or 3 seconds)

<sup>1</sup>“Move here” is available only in the room scenario.

In our current implementation, the robots are remotely controlled by a computer. This computer receives a copy of the instructions from the augmented reality device and combines it with the current positions of objects and robot obtained from overhead cameras. The instructions provided from the augmented reality device are of the following form:

Move\_To(Mug A, Position B, Offset (x,y))  
Operate(Appliance C, Mug A, Action D, Parameters E)

The computer converts these instructions to low-level robot control signals. The robots can move forward, move backward, turn left, turn right, grab, and release using a wireless Bluetooth connection. The process of converting high-level instructions to low-level robot control signals is situated within automated planning and scheduling, a sub-topic of artificial intelligence research. Methods such as STRIPS (Stanford Research Institute Problem Solver) and PDDL (Planning Domain Definition Language) can be applied to this problem, but it is not the main focus of this research.

### C. Home Appliances

We built three custom-made home appliances in our home environment: a pill dispenser, a powder dispenser, and an electric hot water kettle. The pill and powder dispensers are built from scratch. They dispense their contents into a container that is placed beneath the device’s outlet. The hot water kettle is built by adding wireless control to an off-the-shelf electric hot water kettle.

Ideally, all appliances will be controllable by robots, either by manipulating its physical user interface (i.e., knobs, dials, and buttons) or electronically. Our research simulated an electronic interface because we expect an appliance-robot interaction more similar to Star Wars’ R2D2 (digital interface instead of a physical interface). This is a realistic assumption given that there is recent research showing how to adapt home appliances for electronic control [26].

### D. User Tests

We created two scenarios, preparing a drink and organizing furniture in a room, to demonstrate our method of instructing a robot using an augmented reality device.

Table 1 lists the objects that were included in the test environment.

We ran the robot successfully in preparing coffee, pills, and delivering water; and in rearranging floor objects. We also invited seven participants, which were recruited from our research lab but did not have experience with robot programming, to test the teaching interface with these tasks. We showed the robot to participants, but the robot was not interactive during the tests. From these preliminary investigations, participants quickly learned the idea of our system. When teaching a robot to reposition an object, about half of the participants preferred photographing goal states (a single photograph showing the ideal set up once a robot finishes) whereas the other half preferred the “move to” instruction. Participants liked the syntax checks provided by the device, but they were confused by the preprogrammed semantic checks and attributed lock-out conditions to recognition issues.

## VI. DISCUSSION AND CONCLUSIONS

Although new objects can be added easily with photographs, photographs make it difficult to generalize the teaching result. It makes it difficult to refer to photographed objects by their attributes (for example, “the nearest mug” or “a red item”) and function (for example, “an item that separates tea leaves from liquid”). For future work, we plan to add a cartoon-like graphic to identify objects based on attributes.

The proposed technique is demonstrated with robots that push objects around. The user interface can be extended to show a variety of actions such as “grab” (from the side) and “forklift” (lift from underneath) by showing additional actions in Figure 2b. Another extension could allow photographed *areas* to be selected. As an example of vacuuming, the user could photograph the area to clean.

Although an all-in-one robot is possible, robots will more likely cooperate with existing home appliances to accomplish tasks. A robot can be taught how to clean clothes by soaping, rinsing, and wringing, but a washing machine already does this sequential task efficiently. A robot will still be required to place objects in the appropriate home appliances and destinations, a scenario which we have described in this paper.

We have only considered how to teach a single sequential task to a robot. We have not considered how to compose sequential tasks into goals (operate until an environmental state is achieved) and behaviors (hierarchical composition of sequential tasks) [21] using the AR interface. We have not yet considered how to specify control structures such as loops and branches in photographs.

In conclusion, this paper presented an augmented reality system for teaching sequential tasks to a robot. The sequence of annotated photographs serves as a visual representation of the procedure for both a human and robot. We showed the feasibility of the method with a prototype robot system using various household robots and home appliances.

## REFERENCES

- [1] J. Sung, H.I. Christensen, R. Grinter. Sketching the future: Assessing user needs for domestic robots, *RO-MAN 2009*, 153-158.
- [2] T. Miller, A. Exley, and W. Schuler. Elements of a spoken language programming interface for robots. In *HRI 2007*, 231-237.
- [3] T. Winograd. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. *Doctoral thesis*, MIT, 1971.
- [4] B. D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robot Auton. Sys.*, 469-483. 2009.
- [5] Y. Kuniyoshi, M. Inaba, and H. Inoue. Learning by watching: extracting reusable task knowledge from visual observation of human performance. *IEEE Trans. Robot. and Autom.*, v.10, 799-822. 1994.
- [6] M. Pardowitz, R. Zöllner and R. Dillmann. Incremental Learning of Task Sequences with Information-Theoretic Metrics. *European Robotics Symposium*, v. 22, 51-63, 2006.
- [7] G. Biggs and B. Macdonald. A survey of robot programming systems. In *Australasian Conf. Robotics and Autom.* 2003.
- [8] S. Calino and A. Billard. Incremental learning of gestures by imitation in a humanoid robot. In *HRI 2007*, 255-262.
- [9] P.T. Cox and T.J. Smedley. Visual Programming for Robot Control. In *IEEE Visual Languages*, 217, 1998.
- [10] H. Lieberman. *Your Wish is My Command: Programming by Example*. Morgan Kaufmann, San Francisco. 2001.
- [11] L. Leifer, M. Van der Loos, D. Lees. Visual language programming: for robot command/control in unstructured environments. In *Conf. on Advanced Robotics*, 31-36. 1991.
- [12] Y. Sugiura, D. Sakamoto, A. Withana, M. Inami, and T. Igarashi. Cooking with Robots: Designing a Household System Working in Open Environments. In *CHI 2010*, 2427-2430.
- [13] D. Kurlander and S. Feiner. 1992. A history-based macro by example system. In *UIST '92*, 99-106.
- [14] J. Aleotti, S. Caselli, and M. Reggiani. Leveraging on a virtual environment for robot programming by demonstration. *Robotics and Autonomous Systems*, v. 47, 153-161. 2004.
- [15] Y. Sugiura, T. Igarashi, H. Takahashi, T.A. Gowon, C. Fernando, M. Sugimoto, and M. Inami. Graphical instruction for a garment folding robot. In *SIGGRAPH 2009 Emerging Technologies*. 2009.
- [16] Alice: rapid prototyping for virtual reality. *Computer Graphics and Applications, IEEE*, vol.15, no.3, pp.8-11, 1995
- [17] M. Tani, K. Yamaashi, K. Tanikoshi, M. Futakawa, and S. Tanifuji. Object-oriented video: interaction with real-world objects through live video. In *CHI '92*, 593-598. 1992.
- [18] D. Sakamoto, K. Honda, M. Inami, and T. Igarashi. Sketch and Run: A Stroke-based Interface for Home Robots. In *CHI 2009*, 197-200.
- [19] K. Liu, D. Sakamoto, M. Inami, and T. Igarashi. Roboshop: multi-layered sketching interface for robot housework assignment and management. In *CHI 2011*, 647-656.
- [20] S. Hashimoto, A. Ostanin, M. Inami, and T. Igarashi. Photograph-Based Interaction for Teaching Object Delivery Tasks to Robots. In *HRI 2010 Late-Breaking*, 153-154.
- [21] J. Saunders, C. L. Nehaniv, and K. Dautenhahn. Teaching robots by moulding behavior and scaffolding the environment, *HRI 06*, 118-125.
- [22] G. Suzuki, S. Aoki, T. Iwamoto, D. Maruyama, T. Koda, N. Kohtake, K. Takashio, H. Tokuda. u-Photo: Interacting with Pervasive Services Using Digital Still Images. In *Pervasive Computing 2005*.
- [23] K. Ishii, S. Zhao, M. Inami, T. Igarashi, and M. Imai. Designing Laser Gesture Interface for Robot Control. In *INTERACT 2009*, 479-492.
- [24] D. Kragic, M. Bjorkman, H. Christensen, J. Eklundh. Vision for robotic object manipulation in domestic settings, *Robotics and Autonomous Systems*, v.52, i.1, p.85-100, 2005.
- [25] M. Luimula, K. Säskilähti, T. Partala, S. Pieskä, J. Alaspää. Remote navigation of a mobile robot in an RFID-augmented environment. *Personal Ubiquitous Computing*. v.14, i.2, 125-136. 2010.
- [26] K. Fukuchi, T. Igarashi, M. Sugimoto, C. Fernando and M. Inami. Push-pins: Design-by-user Approach to Home Automation Programming, *UBICOMP 2009* workshop.