Paper:

TouchMe: An Augmented Reality Interface for Remote Robot Control

Sunao Hashimoto[*1,*5], **Akihiko Ishida**[*2,*5], **Masahiko Inami**[*3,*5], **and Takeo Igarashi**[*4,*5]

[*1]Meiji University
4-21-1 Nakano, Nakano-ku, Tokyo 164-8525, Japan
E-mail: hashimoto@kougaku-navi.net
[*2]Tokyo University of Science
1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan
E-mail: ishida@designinterface.jp
[*3]Keio University
4-1-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa 223-8521, Japan
E-mail: inami@inami.info
[*4]The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
E-mail: takeo@acm.org
[*5]JST ERATO Igarashi Design Interface Project
7-3-1 Hongou, Bunkyo-ku, Tokyo 113-8656, Japan
[Received October 18, 2012; accepted April 22, 2013]

General remote-control robots are manipulated by joysticks or game pads. These are difficult for inexperienced users, however, because the relationship between user input and the resulting robot movement may not be intuitive, e.g., tilting the joystick to the right to rotate the robot left. To solve this problem, we propose a touch-based interface called TouchMe for controlling a robot remotely from a third-person point of view. This interface allows the user to directly manipulate individual parts of a robot by touching it as seen by a camera. Our system provides intuitive operation allowing the user to use it with minimal training. In this paper, we describe TouchMe interaction and prototype implementation. We also introduce three types of movement for controlling the robot in response to user interaction and report on results of an empirical comparison of these methods.

Keywords: remote robot control, augmented reality, touch screen, direct manipulation, third-person view

1. Introduction

There are many environments in which it is difficult for human beings to work, such as in the ocean, in high places, at very high or low temperature, and in contaminated areas. Various robots have been developed to perform tasks in these dangerous environments and while fully autonomous robot operation is preferable, it is difficult due to recognition problems. Methods for alleviating these problems include putting tags on objects and using prestructured environment models. These methods are not applicable in unstructured environments, however, and human supervisory control is necessary.

A robot that can picks up and delivers physical objects generally has multiple degrees of freedom (DOF), but these are difficult for inexperienced operators to control. The robotic arm generally has 4 or 6 DOF, for example, and total DOF is increased when the robotic arm is mounted on a mobile vehicle. The most common devices for controlling multi-DOF robots are joysticks and game pads; but the number of controllable DOF is limited by the number of buttons and axes of these devices. Controllable DOF is increased by combining two or three keys, for example, but this makes operability more difficult and demands longer user training. Inverse kinematics (IK) is widely used to facilitate the control of a multiple-link robot, and the robotic arm generally uses IK to control the end-effector to a joystick, calculating angles of individual joints appropriately. The velocity of robot movement for the general joystick-based controller is proportional to the timing or degree of key pressing, however, and this also requires user training.

We propose a teleoperation system called TouchMe that allows the user to manipulate a multi-DOF robot intuitively by touch using a third-person point of view. **Fig. 1** shows the touch-screen-based interface, which displays images captured by a camera observing the robot remotely. Users specify desired poses and positions of robots by touching and dragging parts to be controlled. Camera images of the robot are overlaid by computer graphics (CG) models synchronized by user manipulation to help users predict how the robot will move. This augmented reality application applies direct manipulation of a posing tool for CG models, such as Poser, to multi-DOF robots in the real world. Remote-control robots typically use a camera on the robot to provide a first-person view, but we use a camera having a third-person view because

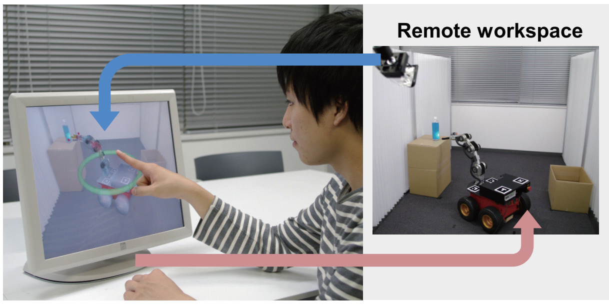

Fig. 1. TouchMe: users directly controlling movable parts of robots by touching camera images.

it helps users better understand the situation in the entire working space – the robot controlled, target objects, and obstacles. We discuss the advantages and disadvantages of cameras with third-person-view settings for the proposed system, i.e., as a fixed-surveillance camera, a flying camera, and a robot-mounted camera.

We describe TouchMe interaction and prototype implementation and introduce three types of movement for robot control in response to user interaction, i.e., moving the robot after touching the interface while touching the interface, and by moving the robot during and after touching the interface. We evaluate these methods empirically and report the results.

2. Related Work

Many video-based interfaces have been proposed for controlling robots and robot applications.

Tani et al. presented interactive video that allows interaction with objects in live video on a screen by having models of objects monitored by cameras [1]. They explored two strategies for modeling objects captured by cameras in 2D and 3D, implementing a system called HyperPlant for monitoring and controlling electrical power plants by using 2D modeling. Seifried et al. developed a video-based interface in the CRISTAL project for controlling home appliances [2]. In this project, the camera is mounted with a bird's-eye view of a living room having images displayed on a multitouch tabletop. This system allows multiple users to operate multiple appliances collaboratively. We also use camera images to control devices, but our controlled object is a multi-DOF robot and we aim to achieve more complicated tasks with it.

A bird's-eye view has been used in several video-based robot control interfaces. Sakamoto et al. proposed a video-based tablet PC interface for controlling vacuum cleaning robots [3]. In this system, ceiling cameras provide users with a bird's-eye view that allows users to control robots and design their behavior by sketching with a stylus. Kato et al. developed a multitouch tabletop interface for controlling multiple robots [4]. They proposed a method of controlling multiple mobile robots simultaneously by manipulating a vector field on a bird's-eye view from a ceiling camera. Guo et al. presented two interfaces for remotely interacting with multiple robots using toys on a large tabletop display showing a bird's-eye view of the workspace [5]. This research shows that a bird's-eye view is easy in controlling the locomotion of mobile robots on a 2D surface; but difficult to control multi-DOF robots.

There are several interfaces for controlling robots through a first-person (robot's-eye view). Sekimoto et al. proposed a simple driving interface for a mobile robot using a touch panel and first-person images from the robot [6]. Once operators set a temporary goal position by touching the monitor displaying the front view of the robot, the system generates a path to a goal and the vehicle follows the path to reach the goal autonomously. Fong et al. developed a similar system on a handheld device (PDA) [7]. Correa et al. proposed a handheld tablet interface for operating autonomous forklifts in which users provide high-level directives to forklifts through a combination of spoken utterances and sketched gestures on the robot's-eye view on the interface [8].

Third-person views are also used in video-based remote robot control. Hosoi et al. proposed robot control called Shepherd that involves a camera-mounted mobile device such as a PDA and a mobile phone [9]. Using this technique, operators hold camera-mounted mobile devices in the hand and instruct the robot how to move by moving the device. Sugimoto et al. proposed visual presentation called Time Follower's Vision for remotely controlling a robotic vehicle [10]. This allows operators to control a remote rescue robot by observing a virtual third-person view from a first-person view by a camera on the robot. They show the effectiveness of the third-person view in allowing even inexperienced operators to control a robot easily.

Several robot interfaces use augmented reality and mixed reality. Nawab et al. proposed overlaying color-coded coordinates on a robot end-effector by using augmented reality to help users understand key assignment of a joystick [11]. Kobayashi et al. developed a mixed reality environment that overlays internal humanoid robot status such as recognition and planning results [12]. This enables operators to understand internal robot status intuitively, which is helpful in debugging and actual operation. Chen et al. developed a mixed reality environment for conducting robot simulation involving physical and virtual objects [13]. Drascic et al. developed an augmented reality application through graphic overlaid on stereo video [14]. In their application, users wearing a data glove control a robotic arm by manipulating a virtual cursor on a video image. Xiong et al. developed a telerobotic system based on augmented reality for controlling a six-DOF robotic arm [15]. A virtual robot works as an interface between operators and real robots, solving the time-delay problem between user operation and robot action. This is also used in our research, but we use a touch screen for the interface and compare three touch interaction methods empirically while wearing a head-mounted display and a data glove and voicing commands for the interface.