Shepherd: A Mobile Interface for Robot Control from a User's Viewpoint

Kazuhiro Hosoi *Masanori Sugimoto*
Interaction Technology Laboratory, Department of Frontier Informatics
Graduate School of Frontier Sciences, University of Tokyo
5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8561, Japan
Email: {hosoi, sugi}@itl.t.u-tokyo.ac.jp
http://www.itl.t.u-tokyo.ac.jp/

ABSTRACT

In this paper, we propose a system called Shepherd that allows a user to manipulate robots such as radio controlled cars from his viewpoint. A user of Shepherd can intuitively control the robots by capturing images of robots through a camera mounted on a mobile device, and moving it. User studies of Shepherd indicated that users could manipulate a robot in an intuitive manner.

ACM Classification H5.2 [Information interfaces and presentation]: Interaction styles.

General Terms Human Factors

KEYWORDS: robot control, mobile interface, user's viewpoint

INTRODUCTION

One problem in remote robot control is related to the difference between the human's viewpoint and the robot's viewpoint. For example, when a user wants to move a robot to the left from the user's viewpoint, he has to send an operation command considering the relative position and direction between the robot and himself. This requires the user to conduct unintuitive operation tasks, which raise the level of his cognitive load, and thus may make him send wrong operation commands to the robot.

Many researchers have proposed human-robot interaction systems (e.g. [1]) by using voices or gestures. In these systems, various sensors are attached to the robot which recognizes the user's gesture and speech, and tries to comprehend his intension. However, such a "sensor-rich" robot is usually expensive, and its intelligent functions are not always necessary, if the task given to the robot is simple (e.g. move to a specified position) and an intuitive technique for the user to manipulate it is provided. Moreover, recognizing voice commands in a noisy environment often becomes difficult and stable control of the robot is not always achieved.

UIST'06, October 15–18, 2006, Montreux, Switzerland.

Another method to solve the problem is to design a sensor-embedded environment where positions of users and robots are automatically recognized [2]. This method may make it possible to control robots without adding any sensors to them. Instead, numerous sensors must be carefully embedded into the environment so that robots and users are always recognized successfully.

In this paper, we discuss an intuitive remote control technique for robots using a mobile device such as a PDA or a mobile phone. A camera-mounted mobile device recognizes the state of the robot and the user's gesture, and makes the robot move as the user has requested. This technique is less expensive than the other methods described above, because almost no modifications to the robot and its surrounding environment are needed. Furthermore, if a camera-mounted mobile phone, which has already penetrated into our society, is available for this technique, it can be used by anyone and anywhere without any special sensors embedded into the robot or its environment.

The proposed system called Shepherd provides the user with a method for intuitively controlling robots from his viewpoint in face-to-face situations. By capturing images of robots by a camera-mounted mobile device and moving it, the user can control them as shown in Figure 2. Shepherd allows the user to control robots without considering the differences between his and their viewpoints.

IMPLEMENTATION

System Overview

Shepherd's system overview is shown in Figure 1. For developing a prototype version of Shepherd, we decided to use mobile PC (Vaio type U by Sony) as it has more computational power compared to PDAs or mobile phones. The USB camera (Qcam Pro4000 by Logicool Inc.) is attached to the mobile PC. The remote controller which is also attached to the mobile PC sends control commands to the robots. The remote controller is connected to the mobile PC via a serial converter. The software of Shepherd has been implemented in Microsoft Visual C++. Shepherd captures through the camera a 320 x 240 pixel image at 20 fps.

In this study, commercially available radio-control cars (about 150 US dollars each) have been used. The size of each is about 0.15 meters (width) $\times$ 0.25 meters (length) $\times$ 0.10 me-

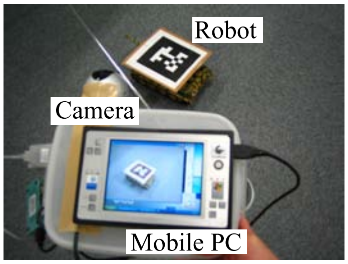

Figure 1: A prototype version of Shepherd

ters (height). It moves around 0.20 meters per second. The robot has two wheels which are separately controlled. For easier recognition of the robot, different visual markers were attached to each individual robot. Shepherd identifies robots' positions and orientations by recognizing the marker. We use the ARTag, an improved version of ARToolKit[3] for recognizing the 3D position and orientation of the robot through its attached maker.

Gesture Recognition

As shown in Figure 2, we assume that the gesture complies with the following rules: 1) When the user holds the mobile PC still, the robot should be at the center of image. 2) When the user moves his mobile PC, this should be done at faster rate than speed of the robot.

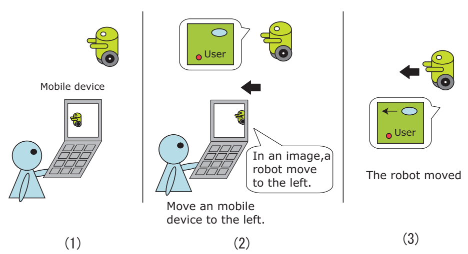

Figure 2: A example usage of Shepherd

Shepherd estimates user's gestures based on the marker's position in a series of images captured by the camera. The following four cases were examined: a) The user is swinging his mobile PC while the robot is moving. b) The user is swinging his mobile PC while the robot is standing still. c) The user holds his PC steadily and the robot is moving. d) The user holds his PC steadily and the robot is not moving. Shepherd calculates the velocity of the marker's motion. Assuming the velocity relation in these cases: $a \simeq b > c > d$, Shepherd can determine the gesture. Informal experiments in these cases proved that a trajectory of the user's gesture could be estimated with sufficient accuracy (recognition rate is 81% at 2 meters away from a robot).

EVALUATION

This paper discusses the effectiveness of user's viewpoint in controlling robots based on a user study. Six university students participated in this study. An experimental setting as shown in Figure 3 was used. The robot is placed at the center of a circle (0.5 meters in radius) directed in one out of 8 possible directions. We selected a random number between 1 and 8 and asked the participants to move the robot in the direction of the number until it is out of the circle. We measured the time between indicating and finishing the task. We

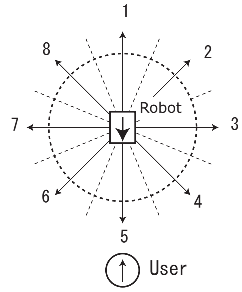

Figure 3: Experimental setting

Table 1: A result of one participant for operation times (sec.). The number of the way means start direction and goal direction.

Remote Controller		Shepherd	
Way	time	Way	time
4 → 1	7.34	6 → 5	4.38
5 → 7	18.47	7 → 8	3.69
8 → 6	6.28	8 → 2	3.75
5 → 4	8.69	3 → 6	3.93
2 → 4	4.37	5 → 4	3.65
avarage	9.03		3.88

repeated the experiment 5 times. They also performed the task using a traditional remote controller that has 4 buttons (left forward, left backward right forward, right backward) in order to compare it with Shepherd. This user study ignored failure of gesture recognition in order to segregate the gesture recognition issues and the viewpoint issues.

The results showed that the average time of operation with Shepherd was 3.80 seconds. Using traditional remote controller, the task took an average of 7.35 seconds to be completed. According to our observation, there was a little difference in the moving time with both controllers. However, using traditional remote controllers, each participant took considerably more operation time when the robot was placed facing the participant himself as shown in Table 1. We also found that participants made the more considerations before moving the robot and they often mishandled the situation, when using the traditional controller.

CONCULSION

In this paper, we proposed a system called Shepherd that allows a user to intuitively control robots from his viewpoint by capturing images of robots with a camera mounted onto a mobile device and moving it within a three dimensional space. User study of Shepherd indicated that it allowed the user to move the robot in an intuitive manner.

At the time of writing, improvements in the robot's control system and gesture recognition are being done. We will further proceed to investigate how to control heterogeneous robots by multi-users.

REFERENCES

1. H. Asoh et.al. Jijo-2: An office robot that communicates and learns. *IEEE Intelligent Systems*, 16(5):46–55, 2001.

2. J. H. Do et.al. Soft remote control system in the intelligent sweet home. In *Proc. of IROS '05*, pages 3984 – 3989, Edmonton, Alberta, Canada, August 2005.

3. H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proc. of IWAR 99*, pages 85–94, San Francisco, USA., 1999.