Orthographic Vision-based Interface with Motion-tracking System for Robot Arm Teleoperation: A Comparative Study

Maram Sakr*
maram.sakr@ubc.ca
University of British Columbia
Vancouver, Canada

Waleed Uddin*
mwaleed16@hotmail.com
University of British Columbia
Vancouver, Canada

H. F. Machiel Van der Loos
vdl@mech.ubc.ca
University of British Columbia
Vancouver, Canada

Abstract

Robot teleoperation is crucial for many hazardous situations such as handling radioactive materials, undersea exploration and firefighting. Visual feedback is essential to increase the operator's situation awareness and thus accurately teleoperate a robot. In addition, the control interface is equally important as the visual feedback for effective teleoperation. In this paper, we propose a simple and cost-effective orthographic visual interface for the teleoperation system by visualizing the remote environment to provide depth information using only a single inexpensive webcam. Further, we provide a simple modification to the control interface (Leap Motion) to achieve a wider workspace and make it more convenient for the user. To realize the merits of the proposed system, a comparison between the modified Leap Motion interface and traditional control modalities (i.e., joystick and keyboard) is conducted using both the proposed orthographic vision system and a traditional binocular vision system. We conduct a user study (N = 10) to evaluate the effectiveness of this approach to teleoperate a 6-DoF arm robot to carry out a pick and place task. The results show that the combination of Leap Motion with the orthographic visual system outperforms all other combinations.

Keywords

Robot Teleoperation; Interface Design; Control Modalities

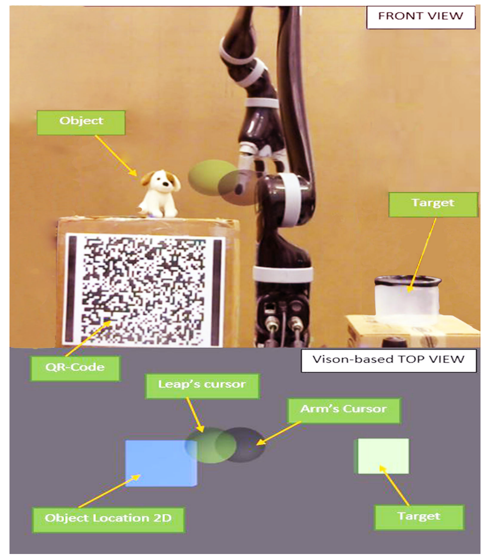

Figure 1: Orthographic visualization top view of the remote location and the camera front view

1 Introduction

Robot teleoperation, also called telerobotics, refers to operating a robot at a distance by a human operator. The control signals are sent from the operator to the robot to control its movement and other signals are sent back telling the operator the current state of the robot. Teleoperation combines human intelligence and maneuverability with robot power that enables the execution of difficult and dangerous tasks without endangering the operator. Over the years, the use of teleoperation has become popular in several areas such as military [5], space exploration [7], mining [2], underwater exploration [12] and telesurgery [8].

Human performance issues related to teleoperating a robot can be classified into two categories: remote perception and remote manipulation [13]. Simple tasks could be challenging due to the lack of visual feedback of the remote environment as well as the mismatching of viewpoints [14][10]. Thus, it is not surprising that several studies found that visual feedback is the most significant contributor to teleoperation task performance [11][3]. In addition, the complexity of the control interface is an essential performance parameter in the operation of the robot [4][6]. Thus, an informative and simple visual interface and an intuitive control mechanism are crucial elements in the teleoperation system [1].

In this paper, we provide a simple, cost-effective and intuitive teleoperation system by visualizing the remote environment in an effective way to provide depth information using only one inexpensive webcam. In addition, a clutch-like system is added to the motion tracking system (Leap Motion [9]) to have a wider workspace and allow the user's hand to move freely above the sensor. To evaluate the benefits of the proposed system, a comparison among three control modalities,joystick, keyboard and the Leap Motion, is conducted using both an orthographic vision system and a traditional binocular system. Thus, our aim is to find the best combination of these modalities and vision systems to result in the fastest and most accurate teleoperation system.

2 SYSTEM DESCRIPTION

We propose a direct control method for a 6-DoF robot in real-time. A 6-DoF Kinova Jaco 2 arm robot is used to evaluate our teleoperation system in the context of a pick and place task. The object to be picked is placed on top of a box that has a quick response (QR) code to help in capturing its position with respect to the camera. The target position has a container into which the operator should drop the object. Two cameras are used to provide visual feedback of the remote location to the teleoperator. The cameras

*These authors contributed equally.

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-7057-8/20/03.
https://doi.org/10.1145/3371382.3378311

are fixed at a distance such that their field of view captures the robot and its surrounding environment.

The operator can control the arm using one of three modalities as follows: (1) Joystick, (2) keyboard, and (3) The Leap Motion sensor (hand motion tracking device). We used the joystick that is available from the manufacturer of the Jaco arm to control the robot. To control the robot using the keyboard, the operator presses the W key to move the robot's end-effector forward, backward with the S key, left and right with the A and D keys, and up and down with the E and Q keys. The space key is to close the gripper, while the X key opens the gripper. For the Leap Motion sensor, it captures and transmits the current hand position as coordinates in X, Y and Z to the computer. A program script handles the incoming data by calculating the absolute distance of the hand movement, then filtering and scaling the data before transmitting to the arm robot through its API over a separate USB link. To control the gripper, we use the *open* and *fist* hand gestures, which are detected by the Leap Motion to open and close the robot gripper. In addition, to solve the limited workspace issue related to the Leap Motion, we implement a clutch concept: the user presses the keyboard space bar to couple/decouple the hand and the robot motion in the case that the hand moves outside the workspace or the user needs a rest. In addition, this avoids any sudden movements by the robot when the hand moves in and out of the Leap Motion workspace.

We tested the teleoperation system with two visual feedback modes as follows: (1) Standard Visual System, and (2) Proposed Orthographic Visual System In the first mode, two cameras are used to provide both the front and top views of the remote environment. In the proposed mode, an orthographic visualization of the remote location is created, as shown in Fig. 1 using only the front view. The blue square represents the object's location with respect to the robot. The green square represents the target position. The black ball represents the robot gripper position in real-time. The green ball represents the operator's hand position if the Leap Motion device is the control modality in use.

3 USER STUDY

In order to assess the overall efficacy of the proposed orthographic vision-based system, we compared it against a conventional vision system while teleoperating a robot arm with the three control modalities. We recruited 10 participants who had no prior experience with robot teleoperation. A consent form approved by the university Behavioural Research Ethics Board was read and signed by each participant prior to the experiment. The participants were asked to complete a pick and place task that consists of picking up a small toy in a gentle way, moving it towards the target position and then dropping it into a target container. The toy is originally placed on an empty cardboard box that can be indented if the participant presses hard on it while picking up the object. This allows us to judge if the participants pick up the object gently.

Participants were introduced to the experimental setup via a verbal briefing. Then, they were provided five minutes to familiarize themselves with the system. Using the standard vision system and our proposed system, participants were asked to complete the pick and place task using Joystick, Leap Motion controller and Keyboard five times each, for a total of 30 trials in random order to permit comparison between the three modalities.

To evaluate our system, we used the task completion time and number of errors as the performance metrics. Errors are defined as the cases in which the participant pressed the box underneath the object (rough pickup); the attempts to pick up the object if the participant did not successfully pick it up the first time; and cases in which the participant did not drop the object in its target position. The number of these errors is counted from the recorded videos of all the experimental trials.

4 RESULTS

The experimental vision system outperforms the standard system regardless of the input modality used, showing a significant reduction in task completion times of 50%, 44%, and 35% using the joystick, keyboard, and Leap Motion, respectively. In addition, the experimental system with the Leap Motion shows the lowest completion time with an average of 37 ± 19 s compared to 42 ± 18 s using the keyboard and 51 ± 23 s using the joystick. For the number of errors, it is comparable among the three modalities, with the experimental system showing higher overall performance than the standard one. The experimental system achieves an average number of errors around 1 ± 1 for all modalities, compared to 3 ± 3, 2 ± 2, and 2 ± 2 using the joystick, keyboard, and Leap Motion, respectively, with the standard vision system.

The two-way ANOVA shows a significant effect of the control modalities ($F(2, 258) = 12.92, p < 0.0001$) and the visual interfaces ($F(1, 258) = 51.45, p < 0.0001$) on task completion time. Similarly, their interaction is statistically significant ($F(2, 258) = 3.75, p < 0.05$). A Bonferroni posthoc pairwise test for the control modalities effect on the completion time shows that using the joystick significantly increases the task completion time more than the keyboard and Leap Motion ($p < 0.05$). Besides, there is no statistical difference between the usage of keyboard and the Leap Motion ($p > 0.05$). In addition, the posthoc test for the effect of the visual interface on the same metric shows that both the standard vision system and the experimental system are statistically different ($p < 0.0001$). Similarly, the two-way ANOVA analysis shows a statistically significant effect of the control modalities ($F(2, 258) = 3.09, p < 0.05$), visual interfaces ($F(1, 258) = 68.58, p < 0.0001$) and their interactions ($F(2, 258) = 4.87, p < 0.05$) to the number of errors metric. The Bonferroni post-hoc pairwise test shows the same pattern as the one with the completion time for both control modalities and visual interfaces.

5 CONCLUSION

In this paper, we proposed an orthographic vision-based system to solve the depth perception problem in teleoperation systems using one inexpensive camera and a QR code reference. In addition, we provided a simple modification to the Leap Motion interface. A comparative study among three control modalities, joystick, keyboard and Leap Motion, using our proposed vision system and a standard binocular camera system was conducted. We evaluated our system through the teleoperation of a 6-DoF robot to do a 3-DoF 'pick and place' task. The performance metrics used were the task completion time and the number of errors. Results show that the usage of the combination of the orthographic vision system and Leap Motion achieved the balance between the speed and accuracy of performing the task compared to the other combinations.

References

[1] Jessie YC Chen, Ellen C Haas, and Michael J Barnes. 2007. Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 6 (2007), 1231–1245.

[2] David W Hainsworth. 2001. Teleoperation user interfaces for mining robotics. *Autonomous Robots* 11, 1 (2001), 19–28.

[3] Ayoung Hong, Heinrich H Bülthoff, and Hyoung Il Son. 2013. A visual and force feedback for multi-robot teleoperation in outdoor environments: A preliminary result. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on.* IEEE, 1471–1478.

[4] David Kent, Carl Saldanha, and Sonia Chernova. 2019. Leveraging depth data in remote robot teleoperation interfaces for general object manipulation. *The International Journal of Robotics Research* (2019), 0278364919888565.

[5] Tomáš Kot and Petr Novák. 2018. Application of virtual reality in teleoperation of the military mobile robotic system TAROS. *International Journal of Advanced Robotic Systems* 15, 1 (2018), 1729881417751545.

[6] Corinna E Lathan and Michael Tracey. 2002. The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence: Teleoperators & virtual environments* 11, 4 (2002), 368–377.

[7] Guanyang Liu, Xuda Geng, Lingzhi Liu, and Yan Wang. 2018. Haptic Based Teleoperation with Master-slave Motion Mapping and Haptic Rendering for Space Exploration. *Chinese Journal of Aeronautics* (2018).

[8] Jacques Marescaux, Joel Leroy, Michel Gagner, Francesco Rubino, Didier Mutter, Michel Vix, Steven E Butner, and Michelle K Smith. 2001. Transatlantic robot-assisted telesurgery. *Nature* 413, 6854 (2001), 379.

[9] Leap Motion. 2015. Leap motion controller. *URl: https://www. leapmotion. com* (2015).

[10] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2019. Remote Telemanipulation with Adapting Viewpoints in Visually Complex Environments. In *Proceedings of the Robotics: Science and Systems.*

[11] A Reiter, V Nitsch, G Reinhart, and B Färber. 2009. Effects of visual and haptic feedback on telepresent micro assembly tasks. In *3rd International Conference on Changeable, Agile, Reconfigurable and Virtual Production CARV.*

[12] Roque Saltaren, Rafael Aracil, Cesar Alvarez, Eugenio Yime, and Jose Maria Sabater. 2007. Field and service applications-Exploring deep sea by teleoperated robot-An Underwater Parallel Robot with High Navigation Capabilities. *IEEE Robotics & Automation Magazine* 14, 3 (2007), 65–75.

[13] Aaron Steinfeld, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. 2006. Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction.* ACM, 33–40.

[14] James S Tittle, Axel Roesler, and David D Woods. 2002. The remote perception problem. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 46. SAGE Publications Sage CA: Los Angeles, CA, 260–264.