Applied Ergonomics xxx (2016) 1–8

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Comparison of three different techniques for camera and motion control of a teleoperated robot

Guillaume Doisy*, Adi Ronen [1], Yael Edan

Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva, Israel

ARTICLE INFO

Article history:
Received 18 August 2014
Received in revised form
16 February 2016
Accepted 1 May 2016
Available online xxx

Keywords:
Teleoperation
Human factors
Interface design

ABSTRACT

This research aims to evaluate new methods for robot motion control and camera orientation control through the operator's head orientation in robot teleoperation tasks. Specifically, the use of head-tracking in a non-invasive way, without immersive virtual reality devices was combined and compared with classical control modes for robot movements and camera control. Three control conditions were tested: 1) a condition with classical joystick control of both the movements of the robot and the robot camera, 2) a condition where the robot movements were controlled by a joystick and the robot camera was controlled by the user head orientation, and 3) a condition where the movements of the robot were controlled by hand gestures and the robot camera was controlled by the user head orientation. Performance, workload metrics and their evolution as the participants gained experience with the system were evaluated in a series of experiments: for each participant, the metrics were recorded during four successive similar trials. Results shows that the concept of robot camera control by user head orientation has the potential of improving the intuitiveness of robot teleoperation interfaces, specifically for novice users. However, more development is needed to reach a margin of progression comparable to a classical joystick interface.

1. Introduction

Remote control of robotic platforms through teleoperation has been used for years in many applications: telepresence (Jouppi, 2002), exploration of remote and hostile locations (Burke et al., 2004), or augmentation of human perception and power (Turro et al., 2001). However, teleoperation is subject to important human factors issues (Chen et al., 2007; Murphy, 2004; Tittle et al., 2002; Voshell et al., 2005). Notably, the experience of the user, the context of use of the robot, the visual information available and the interaction modality have strong effects on the user performance (Casper and Murphy, 2003; Scholtz et al., 2004; Woods et al., 2004). Also, there is a strong influence of the operator's situational awareness (Drury et al., 2003; Yanco and Drury, 2004), the operator's understanding of the robot's close and far surroundings. In a teleoperated system that uses an embedded camera, the "keyhole" or "soda straw" effect is a major factor degrading situational awareness (Casper and Murphy, 2003; Voshell et al., 2005).

Hence, significant effort has been put in finding a solution for the operator to make better sense and use of the robot camera. The limited field of view of the camera has quickly been identified as a major issue (Alfano and Michel, 1990) and to overcome it some studies suggested the use of wide angle lens on the camera (Eliav et al., 2011; Scribner and Gombash, 1998).

However, when receiving the video feed of a camera equipped with a wide angle lens, users have been reported to over-estimate the speed of the objects in their surroundings and therefore to inappropriately reduce their speed (Scribner and Gombash, 1998). Additionally, the optical distortion caused by a wide angle can increase the risk of operator motion sickness as well as the cognitive workload to operate the robot (Chen et al., 2007). Alternatively, it has been suggested to use multiple cameras, like in (Keyes and Yanco, 2006) where a rear-facing camera was added; or in (Voshell et al., 2005) where 5 cameras (one pointing straight and the four others pointing 45° in each direction: up, down, left, right) were used to create *Perspective Folding*. But with multiple cameras

* Corresponding author.
E-mail addresses: doisyg@post.bgu.ac.il (G. Doisy), adiro@bgu.ac.il (A. Ronen), yael@bgu.ac.il (Y. Edan).

[1] In Memoriam: Dr. Ronen passed away during the final stages of preparing this manuscript, over the last few years he has conducted and promoted research in the area of human-robot collaboration, His untimely demise is a great loss to the human factors community.

http://dx.doi.org/10.1016/j.apergo.2016.05.001
0003-6870/

there is a risk of cognitive tunneling (Thomas and Wickens, 2000) that occurs when the operator's attention is captured by a single camera output and the feeds from the other views are ignored. Another suggestion which was widely adopted is to allow the decoupled motion of the camera orientation from the movements of the robot (Hughes et al., 2003); i.e., to mount the robot camera on a pan-tilt (or pan alone) mechanism and to provide the operator with independent control of the mechanism. But this solution presents the risk of degrading further the operator's situational awareness instead of increasing it if not carefully used (Nielsen et al., 2005). It appeared that when using a controllable camera orientation, some users have trouble detecting when the camera is not aligned with the forward direction of the robot which can provoke dramatic collision and loss of the system (Drury et al., 2003). To overcome these problems of unnoticed misalignment of camera orientation two main strategies can be found in the literature.

The first approach, described in (Nielsen et al., 2007; Nielsen et al., 2005; Ricks et al., 2004), is a specific form of ecological design for teleoperation. The problem of misalignment is solved by constructing a 3D virtual exocentric view of the robot. Inside this view the output of the camera can be displayed according to the current orientation of the pan-tilt mechanism. Hence, the operator can directly visualize the position of the camera with respect to the robot, as illustrated in Fig. 1. However, the process required to construct the 3D view is complex to set up and needs reliable localization in addition to an accurate range sensor, limiting its usage for simple robotic hardware. Nevertheless, this approach has proven to provide advantages in terms of performance and operator workload as compared to classical interfaces.

A second approach consists of controlling the orientation of the video feedback through the operator's head orientation: conscious of their head orientation, operators are then aware when there is a misalignment between the forward robot direction and the video feedback. In (Zalud, 2006) for instance, the operator uses virtual reality goggles capable of tracking the head orientation, which is then used to control the pan-tilt mechanism. However, no comparison was made with classical camera orientation control. Moreover, the latency between the head orientation, the actual movement of the pan-tilt mechanism and the update of the video image in the head mounted display is likely to provoke motion sickness, discomfort and degradation of perceptual capabilities typical of virtual reality display with high latency (Allison et al., 2001; Mania et al., 2004). This latency issue was overcome by the use of a similar head tracking virtual reality goggles in conjunction with omnidirectional cameras, like in (Fiala, 2005). The operators can then choose their video view orientation though their head movements with very little delay since there is no need to wait for a mechanical device to move and an image the be transmitted. However, head tracking virtual goggles and see-through techniques are promising but heavy to wear and exhausting for the operator and a significant minority of the population still cannot use them without experiencing motion sickness with the current state of the technology. Additionally, visual displacement caused by see-through systems deteriorates visio-motor performance due to sensory conflict (Biocca and Rolland, 1998; Cobb, 1999; Smyth, 2000) and no user studies proved significant improvement of robot teleoperation performance with such systems compared to classical control modes.

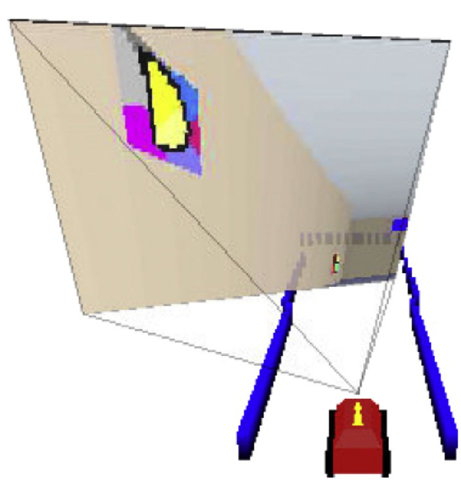

Fig. 1. Visualizing the orientation of the pan-tilt camera using the 3D interface, from (Nielsen et al., 2007). In this specific work, the experiment was conducted in a simulation and hence the video feed is also simulated, which is the reason why it is not photorealistic.

This research aims to evaluate a new method for controlling the orientation of the camera through the operator's head orientation in robot teleoperation tasks. Specifically, the use of head-tracking in a non-invasive way, without immersive virtual reality devices was combined with joystick or hand gesture robot control, and compared in terms of performance and workload to a classical robot teleoperation interface. Additionally, the effect of the user experience and the way performance and workload evolved through consecutive trials was investigated.

Teleoperation performance was evaluated through task completion time and number of collisions with the walls, as it is often done in research comparing different robot teleoperation interfaces, such as in (Nielsen et al., 2007). Heart rate (HR) may be used as an indication of the physiological state of the participants and may be indicative to workload levels, fatigue and physiological strain (Roscoe, 1992, 1993; Turner and Carroll, 1985), such as in (Harriott and Zhang, 2011) where HR measurements were used in a human-robot interaction experiment. Hence, it was used in this research to objectively measure workload. In addition, subjective workload was assessed through the raw NASA-TLX questionnaire (Hart and Staveland, 1988), similarly as in (Nielsen et al., 2007) human-robot experiment.

2. Methodology

Thirty-six industrial engineering students, 21 males and 15 females between 22 and 28 years old, with no previous teleoperation experience were recruited through email. Participants received a compensation of 30 Israeli Shekels (NIS), about 8 United States Dollars (USD), for their participation and were told prior to the experiment that they could potentially win a bonus of 100 NIS (about 26 USD) through roulette wheel selection depending on their performance. The higher the score of a subject was, the more virtual lottery tickets she/he received. Among all the virtual lottery tickets distributed, one was chosen to receive the 100 NIS bonus.

This experiment used a mixed between and within-subject design.

Participants were split into three groups of twelve students each for the three experimental conditions A, B and C which differed in the way the robot camera orientation and the robot movements were controlled, as described in Section 2.3. The group was the between-subject variable.

The exploration task was conducted four times to study the impact of experience on performance. The trial number was the within-subject variable.

Please cite this article in press as: Doisy, G., et al., Comparison of three different techniques for camera and motion control of a teleoperated robot, Applied Ergonomics (2016), http://dx.doi.org/10.1016/j.apergo.2016.05.001