Image Display Virtualization in Teleoperation

Jonathan M. Smith, Brian P. DeJong, Ernur Karadogan, James Hasbany
School of Engineering & Technology
Central Michigan University
Mount Pleasant, MI
smith122@cmich.edu

***Abstract*—For successful teleoperation, an operator must be able to accurately inscribe the desired movement to the manipulandum. The goal of manipulating an image display for teleoperation is to perform visual transformations that minimize mental fatigue, performance time, and enhances the compatibility between the operator and the teleoperated system for successful task completion. A concept based on multiple physical and virtualized displays will identify the areas in which the increasing mental demand reduces the effectiveness of the operator. The effect of physical and virtual display configurations was investigated in an experimental study. A total of 18 participants teleoperated an industrial robot under three display configurations. The results of the study revealed that the virtualization of the display is a more effective way to present information to the operator of a teleoperated system when the goal-to-goal times and the total distance traveled during a task were considered.**

***Index Terms*—teleoperation; mental workload; virtualization**

I. Introduction and Background

Teleoperation allows users to remotely control robots and apparati in environments that could otherwise be inaccessible. Teleoperated systems have increasingly been used in critical domains such as urban search and rescue, exploration tasks (space, deep-sea, volcanoes, etc.), remote surgery, and high-cost industrial inspections (e.g., with nuclear reactors) [1,2]. Many skilled operators are required to fill the increasing demand for robotic teleoperation.

Teleoperation is mentally challenging for users in part because of the limited and often unintuitive visual feedback [3]. The user must use limited video images of the remote environment and robot. If using a joystick (or similar device) to control the robot, the user is often forced to learn the hand-eye mapping between inputs and the visual feedback – e.g., pushing right on the joystick does not necessarily move the robot to the right in all, if any, video images. Such hand-eye alignment has been shown to increase mental workload [4], task time and errors [5]. The more joints of the teleoperated robot, the higher the user's mental workload [6]. While much research involves automating the robots and increasing the information fed back to the user (visual or otherwise), much less attention has been given to simplifying that information or making it more intuitive. Immersing the operator in an augmented reality or using a head mounted device is not always feasible and has even been shown to be ineffective [7].

Previous work has shown that properly aligning video images on separate, moveable displays can significantly improve teleoperation task time and performance [5]. Introducing haptic feedback has also been shown to improve operator performance within the range of 50-100ms [8]. However, many teleoperation interfaces do not or cannot include multiple moveable monitors to allow for proper video alignment and cannot include bulky haptics modules. Robots such as the ones used in the military field require a light simple system of no more than 25lbs [9]. With the limitations of weight and space, maximizing the reliability of the system and minimizing the system weight and dimensions give the operator increased information while reducing the equipment load [10].

Exploring the factors that affect the mental demand an operator must endure is imperative to create a more suitable interface regardless of whether it is an augmented environment with haptic feedback or virtualizing the display on a monitor. In that direction, this study aims to investigate if displaying multiple video images on a single stationary display has similar benefits as physically rotating displays. We hypothesize that, by manipulating the visual feedback provided to the operator, task times and performance can be improved without significantly increasing the mental workload.

The following sections present and discuss a human-subject experiment into whether virtualized displays improve teleoperation. Section 2 discusses system setup and experimental methods. Section 3 presents results. Section 4 discusses the results, conclusions and lessons learned, and future work.

II. Methods

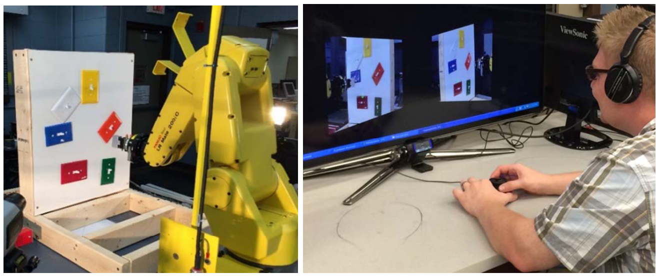
A. Participants

This study included a total of 18 self-selected participants (14 male, 4 female). The average age of the participants was 32.9±11.3 (mean ± S.D.). The participants did not have any previous experience in comprehensive teleoperation.

B. System Setup

The system (shown in Fig. 1) consisted of a computer with proprietary user interface, an input device, and one or more monitors depending on the interface being tested. The users viewed the teleoperated robot via two video images displayed and distorted following the experimental protocol. In addition, the user wore noise canceling headphones to minimize

978-1-5090-3202-0/17/$31.00 ©2017 IEEE



(a) (b)

Fig. 1. Teleoperation setup consisting of (a) the user interface and (b) remote robot and environment.

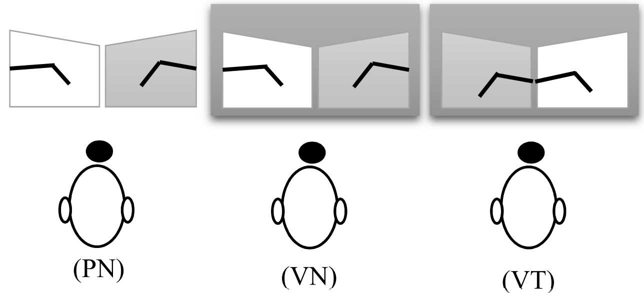


Fig. 2. The three setups tested in the experiment: Physical Normal (PN), Virtualized Normal (VN), and Virtualized Transposed (VT).

distraction and auditory cues from the robot; the user could not see or hear the robot directly. The input device used by the participants was a six-degree-of-freedom 3Dconnexion Space Navigator which controlled the robot's end effector in world coordinates. The computer read the user's input, commanded motions of the robot via direct connection to its controller, received position and video feedback, and distorted and displayed the video images on the monitor(s).

The robot, task environment, and video cameras were arranged remotely from the user. The teleoperated robot was a six-degree-of-freedom Fanuc LR Mate 200 iC with attached (deactivated) gripper. The robot had integrated force sensing and automatically stopped during collisions. Two CCTV video cameras were mounted on tripods and connected via a frame grabber video card to the interface computer. In addition, two industrial floodlights were used to illuminate the task environment; the resulting faint shadow allowed the user to estimate the robot's distance from the objects.

C. Experimental Procedure

This study tested three video image setups, shown in Fig. 2. The first setup was a Physical Normal (PN) arrangement consisting of video images on two separate display that were physically rotated for proper hand eye alignment (using the technique discussed in [11]). Thus, pushing right on the joystick, for example, moved both robot images to the right as seen by the user. The second and third setups displayed the video images on a single, larger display but at the same pixel and physical size. The images were distorted to mimic the physical rotation of the separate displays, either in a pseudo-properly-aligned (Virtual Normal; VN) or improperly-aligned (Virtual Transposed; VT) arrangement. In the VT setup, pushing right on the joystick, for example, moved the robot images in misaligned directions.

Virtualizing the images does *not* generate the same visualization of the task scene as physically rotating the displays – the "distant" part of the images are distorted with some loss of pixel information, and the alignment does not change with user movement (as it does with the PN setup). However, these setups represent the common use of a single monitor in many teleoperation applications. Thus, the three setups allow for two useful comparisons:

(1) Comparison between physical and virtualized displays (PN versus VN).
(2) Comparison between virtualized aligned and misaligned displays (VN versus VT).

The task environment consisted of six standard light switches arranged on a vertical wall (see Fig. 1.a). The backside of the switchboard was weighted down to reduce the movement from the robotic arm encountering the switchboard surface. The user's task was to activate each switch using the robot's end effector as prompted by the computer. For each test, the switches on the board began oriented towards the center of the workspace. The user was then prompted to move to a specific switch designated by a green augmented reality arrow on the display. They would continue to move and follow and switch the designated switches.

For the experiment, subjects were briefly introduced to the components and protocol. They were not informed about the specific goal of the test. They were given an initial survey on their teleoperation and gaming experience.

Each participant completed three tests with different interface setups and task sequences. The presentation order was pseudo-randomized for each participant for no repeats. At the start of each test, participants were allowed up to five minutes to familiarize themselves with the control and video feedback. Once satisfied, the participant would then manipulate the robotic arm and complete the tasks. The computer recorded and timed user input and position feedback for post processing. After each test, the participant completed a Setup Survey including a NASA TLX mental workload questionnaire [12].

Upon completion of all three tests, the participant filled out a final survey asking their overall impressions of the interfaces.

D. Statistical Analyses

The data were analyzed using repeated-measures ANOVA. The dependent variables were the goal-to-goal times, distance traveled, and the variables obtained from the NASA TLX mental workload questionnaire (mental demand, physical demand, temporal demand, performance, effort, and frustration level). The ANOVA was preceded by a Mauchly's sphericity test to check whether the variances were homogeneous. If this test was significant ($p < .05$)—indicating a violation of the assumption of homogeneity of variance—Greenhouse-Geisser correction procedure was used to adjust the degrees of freedom. Follow-up post hoc comparisons (LSD) were applied when necessary to identify statistically significant changes. A three-sigma test was performed on all data before analysis for outlier detection. The effect of the setup order (between PN, VN, and VT) on each dependent variable was initially tested and no significant order effect was reported.

III. Results

A. Surveys

The initial survey results showed that the average participant had 12 years of video game experience ranging from 0 to 35 years. Participants with experience on consoles made up 85% of all subjects. The other participants had their video game experience on a PC or both. 13% of the subjects did not have any video experience at all. Thirty percent of the participants had some type of Remote Control (RC) experience — one participant had experience related to the control of an RC robot.

The repeated-measures ANOVA revealed no significant difference across the setups when the variables (mental demand, physical demand, temporal demand, performance, effort, and frustration level) obtained from the NASA TLX mental workload questionnaire are considered. The results of the initial and final surveys indicated that there was not a clear bias towards one setup or another. The participants found all the screen orientations to feel similar to them and in the case of the virtualized display, most did not realize what change had been made to the display orientation.

In the final surveys the participants did not indicate any clear preference to one setup over another overall. Most subjects mentioned in their surveys that they felt more comfortable with their final setup regardless of the setup (PN, VN, or VT). Despite what the analysis of the data shows, most participants reported that they performed the best on the PN setup, only one participant reported better performance on the VN setup.

B. Test and Goal-to-Goal Times

The average total test times (time to flip all switches) are shown in Fig. 3. There was no significant difference between the test times across the setups ($F_{(2,34)} = 1.560$). They show that there were significant reductions in time when performing on a virtualized display over a physical display. The total time it took a participant to perform a setup did not show any significant difference regardless of the display being virtualized or physical. There was also no significant difference based on operators' video game experience.

The Goal-to-Goal times for each setup (time from one switch to the next) are shown in Fig. 4. Results show significant differences between the Physical Normal (PN) and Virtual Normal (VN) as well as Virtual Transposed (VT) ($p < 0.05$ for both cases) setups. There was no significant difference between the setups VN and VT.

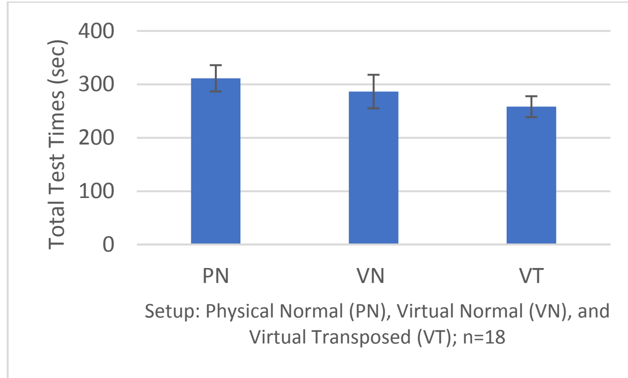


Fig. 3. Average total test times for the tested setups (mean ± standard error)

C. Distances

The average distance traveled for each setup is shown in Fig. 5. There was a significant effect of setup on the distance traveled ($F_{(2,26)} = 5.356$, $p<0.05$, $\eta^2 = 0.292$). The follow up post hoc comparisons revealed that the participants traveled significantly more during the Physical Normal (PN) setup as compared to both Virtual Normal and Virtual Transposed setups ($p<0.05$ for both cases). There was no significant difference between the setups VN and VT. The different total distances traveled in each

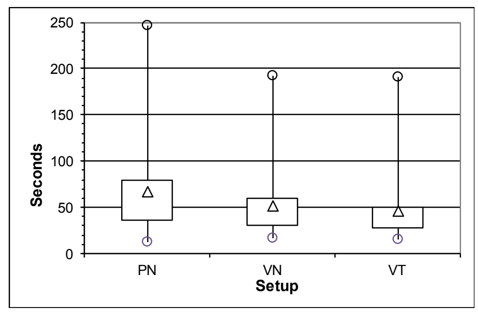


Fig. 4. Quartile plot showing the distribution of goal-to-goal times for the three setups.

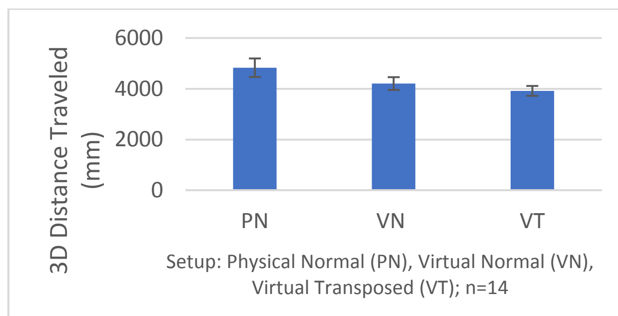


Fig. 5. Average three-dimensional distance traveled for the tested setups (mean ± standard error)

setup show a significant bias towards a virtualized setup. Both the VN and VT setups were similar in results and showed no significance in the difference of the distances traveled during those setups for the participants. In addition, participants' video game experience had no significance on the distances.

IV. Discussion and Conclusions

The results indicate that the virtualized screens perform better than multiple physical screens. This may be due to the complexity of the procedure, the time and fatigue caused by working out three separate setups and finally the apparent ease of interacting with a virtual setup.

Anecdotally, there was a large difference in the participant's bodily movement between physical and virtualized setups. The virtualized displays had almost no movement, however when participants had multiple physical displays they would shift their bodies side to side, and forward. The physical displays were oriented so that a participant had the freedom to move in and out of the image plane, which they did frequently during this setup. They also found themselves moving their heads and bodies around to gain a better perspective outside of the image plane. This was not the case when the images were virtualized. The participants remained mostly still, not being able to move through the image plane. This could increase the focus on both images.

Interacting with the flat surface of the switches was a relatively straightforward task for the operator to complete. In the future, a more complicated task might help distinguish between setups. Forcing the participant to interact with multiple challenges that vary not only in a single plane, but also into a space might provide more useful results. Since it was shown that it does not take a significantly different amount of mental effort to perform this task with the virtualized displays inverted, removing that from the next study would save time and open the direct comparison between physical displays and a virtualized system of displays.

The use of two virtualized displays on a single monitor seems to be the more effective way for an operator to interface with the system. This is the most ideal outcome for this study as it implies that the reduction in information when simulating a display does not increase mental workload, task time, or task performance. A virtualized setup – even a misaligned one – may be easier for an untrained operator to use than a physical one. However further research is needed to fully understand if the results will carry over and whether a trained operator would see the benefits while using virtualized displays.

In the future, it might be beneficial to select operators to represent a sampling of typical teleoperators, for more applicable results. Furthermore, a longer experiment into mental fatigue may enhance the distinction between screen orientations.

Acknowledgment

This work was funded in part by the National Science Foundation's Research Experience for Teachers grant #EEC-1542368. The authors thank the additional contributions of researchers Sandra Dubridge, William Gustin, and Ariel Sanchez.

References

[1] A. Singh, S.H. Seo, Y. Hashish, M. Nakane, J.E. Young and A. Bunt, "An interface for remote robotic manipulator control that reduces task load and fatigue," 2013 IEEE RO-MAN, Gyeongju, 2013, pp. 738-743.

[2] A. Dąbrowska, M.B. Jaskółowski and A. Rubiec, "Cameras vibrations influence on efficiency of teleoperated Unmanned Ground Vehicle," 21st International Conference on Methods and Models in Automation and Robotics, Miedzyzdroje, 2016, pp. 772-777.

[3] J.Y.C. Chen, E.C. Haas, M.J. Barnes, "Human Performance Issues and User Interface Design for Teleoperated Robots," IEEE Transactions on Systems, Man, and Cybernetics, 37:6, 2007, pp. 1231-1245.

[4] D.B. Kaber, J.M. Riley, R. Zhou, J. Draper, "Effects of Visual Interface Design, Control Mode and Latency on Performance, Telepresence and Workload in a Teleoperation Task," Human Factors and Ergonomics Society Annual Meeting, July 2000.

[5] B.P. DeJong, J.E. Colgate and M.A. Peshkin, "Improving teleoperation: reducing mental rotations and translations,", IEEE International Conference on Robotics and Automation, 2004, pp. 3708-3714.

[6] E. S. Neo, K. Yokoi, S. Kajita and K. Tanie, "Whole-Body Motion Generation Integrating Operator's Intention and Robot's Autonomy in Controlling Humanoid Robots," in IEEE Transactions on Robotics, vol. 23, no. 4, pp. 763-775, Aug. 2007.

[7] S. Nahavandi, J. Mullins, M. Fielding, H. Abdi and Z. Najdovski, "Countering Improvised Explosive Devices through a multi-point haptic teleoperation system," IEEE International Symposium on Systems Engineering, Rome, 2015, pp. 190-197.

[8] J. Ramos, A. Wang, S. Kim, "A balance feedback human machine interface for humanoid teleoperation in dynamic tasks", IEEE/RSJ Internation Conference on Itelligent Robots and Systems (IROS), 2015, pp.4229-4235

[9] J.D. Carier, E.D. Smith, A.M. Wade, P.S. Walker and M. J. Kwinn, "Small unit unmanned weapon system for today's army," IEEE Systems and Information Engineering Design Symposium, Charlottesville, 2007. H. Alemzadeh, D. Chen, X. Li, T. Kesavadas, Z.T. Kalbarczyk and R. K. Iyer, "Targeted Attacks on Teleoperated Surgical Robots: Dynamic Model-Based Detection and Mitigation," 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Toulouse, 2016, pp. 395-406.

[10] H. Alemzadeh, D. Chen, X. Li, T. Kesavadas, Z.T. Kalbarczyk and R. K. Iyer, "Targeted Attacks on Teleoperated Surgical Robots: Dynamic Model-Based Detection and Mitigation," 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Toulouse, 2016, pp. 395-406.

[11] B.P. DeJong, *et al.*, "Lessons learned from a novel teleoperation testbed", Industrial Robot: An International Journal, 33:3, 2006, p.187 – 193.

[12] S.G. Hart, “NASA-Task Load Index (NASA-TLX); 20 Years Later,” Human Factors and Ergonomics Society Annual Meeting, October 2006.