

Evaluating Operator Performance in Teleoperated Manipulator System Factored by Camera Configurations

Yan Fu, Wei Chen^(✉), Shiqi Li, and Zhang Jiao

School of Mechanical Science and Engineering, Huazhong University of Science
and Technology, Wuhan 430074, Hubei Province, China
Laura_fy@mail.hust.edu.cn

Abstract. Tele-operated manipulator has been used in a variety of occupational situations, ranging from space exploration, undersea operations, special search and rescue activities to robotic surgery. It is a challenging task, due to the difficulty with maintaining awareness of the robotic arm's configuration and arm location in the task environment. Operators sometimes fail to understand the relative position posture in task scenario as well as arm configuration such as joint limits or singularities due to multi-channel visual display and poor ability in integrating all channels of visual information. The present study served to understand how operators build up situation awareness of location and arm configuration in the task scenario with multi-channels of visual information. The paper mainly discusses the nature of spatial ability with handle controller in time delay condition, and concerns with the development of routines for making task allocation and sequences adaptable. The findings from experiments indicated the best camera configuration and how automated routine allocation does not help too much improve the performance.

Keywords: Space cognition · Teleoperation · Human performance · Camera configuration

1 Introduction

A tele-operator is a machine that enables a human operator to move about, sense and mechanically manipulate objects at a distance. It usually has artificial sensors and effectors for manipulation and or mobility, together with a means for the human to communicate with both. The supervisory part of the tele-operator is positioned in a clean and safe master environment [1]. Based on the received information feedback from the slave system environment, the human operator controls the tele-operator task accomplishment. The master system usually makes use of a sensor which detects human operator movements and transmits them to the slave system device. Joysticks or exoskeleton devices are often used as master system sensors. Teleoperation is the communication with and control of a remotely located machine by a human operator [2].

Human performance issues involved in tele-operating systems generally are two parts: remote perception and remote manipulation. Remote perception refers to perceptual ability of the operator to precisely obtain the information about target location

and orientation to guide moving or crossing an obstacle by visual. In the tele-operating environments, human perception is often inaccurate because the natural perceptual processing is decoupled from the physical environment. This decoupling is derived from time lags, lack of proprioception, frame of reference, and attention switches, affecting people's perception of affordances in the environment [3]. Remote manipulation includes both navigation and manipulation tasks. In order to identify task demands and operator plans for achieving goals and reduce high workload, ergonomic task analysis is used to help explain the impact of stage of automation on human performance in tele-operated manipulator control task, and used for simulating human-machine integrations to optimize manipulation, even simplify some high workload steps under range of condition with human-machine function allocation [4].

It is important to understand the implications of breaking the natural link between human cognitive and perceptual processing and our presence in a physical environment. In the remote perception situation, the observer cannot directly interact with the environment and the perception-action link is broken. Such de-coupling of the natural elements of the environment and the perceptual system means that remote vision hobbles the basic competencies of human perception because it starts with an impoverished stimulus [5]. Tele-operated manipulator tends to extremely difficult to operator, because human performance is "limited by the operator's motor skills and his ability to maintain situation awareness of remote environments, distance estimation and obstacle detection can also be difficult" [6]. As developers use the new powers of technology, they are discovering that it can be surprisingly difficult to understand a physical environment when that environment is only perceived remotely through sensor and video feeds from robotic platforms [7]. Why do these and other ambiguities arise in remote perception when we so fluently and quickly apprehend our environment and what it affords us when we are directly present?

Several factors contribute to difficulties with maintaining awareness of the arm configuration. First and basically, in the remote vision situation the observer cannot directly interact with the environment and the perception-action link is broken. Such de-coupling of the natural elements of the environment and the perceptual system means that remote vision is an example of perception based on an impoverished stimulus. Secondly, more abstract aspects of building a shared perspective (e.g., situation awareness) depend on quick accurate pick up of the affordances of the environment based on perception of high level properties such as point of view, relative scale, and rate of approach to obstacles. For example, the current robotic arm interface shows only the current angle but does not provide reference points, such as the problematic ranges for singularities and joint limits. Even if the design of the joint angle display is improved, another problem with monitoring the arm configuration remains. Operators need to divide their attention between multiple visual tasks and numerous displays/task views that are distributed across several monitors [8]. A possible solution to reduce scanning costs by integrating important arm configuration information with view of the arm and workspace on the monitor. The interface should support this integration through "display proximity" by displaying both pieces of information near each other or even superimposed [9]. Thus, it is worthwhile to work out the best integration view system for efficient visual perception about the arm

position and altitude. Then, the paper mainly focuses on the nature of spatial cognition ability of teleoperators when under different configuration of cameras.

Meanwhile, like any human performance requiring feedback, teleoperation will be less effective if there are time delays or lags in the system. It was found that complex tasks could be accomplished better by adopting a feasible strategy of performing the task by a series of discrete, open-loop movements along with predictive demonstrations to help the human overcome the delay. A classical approach to time-delayed teleoperation requires the operator to specify waypoints which the users send to the manipulator, which drives itself to them. Although effective, few human decisive feedbacks are available to participate in the loop system in this approach. Another reason is that there is a history of problems when the system automation fails. Hence, tele-operated manipulator control options including feedback decision loops under delay conditions indicate that the operator will be in the loop of the system, either as a planner or a controller. This kind system can use adaptive automation algorithms that return control of manipulator systems to the human operator under specified environmental conditions such as a requirement for a tactical maneuver or safe operations [10], it also provides a way of easing workload of operators and adapting their control to change from focusing on manipulation to making reasonable decision, making them attend increasingly to information the system feedback. Although previous research suggested that adaptive automation algorithms, natural interaction and advanced human-machine interfaces that could facilitate telepresence to improve human performance. There are little experimental evidence to substantiate a link between trajectory automation recommended by human-machine function allocation and spatial location and orientation cognition in a complex task condition. So it is another goal of this study to understand how special cognition improves with trajectory automation recommended by virtual reality interface.

2 Method

2.1 Participants

Twelve university students participated in the experiment (6 males and 6 females; mean age = 23.1 SD = 1.2). They reported normal or correct-to-normal vision. None of them had previous experience with operating a robotic arm. Each participant was asked to perform fly-to-tasks, i.e. they had to move the robotic arm to a given target location and orientation. This task involves five steps: (1) selecting 3 camera views; (2) selecting 3 frames of reference to define the location and the orientation of the arm and the relationship between hand controller inputs and the movement of the arm, (3) selecting the rate of arm movement, (4) selecting the control mode (manual mode or supervisory mode), (5) operating the arm with two hand controllers. The translational hand controller controls the movement along X, Y, Z axis; the rotational hand controller controls rotations about 3 axes. While operating the arm, the operators are to exert multiple subtasks including (1) Recall the plan for hand controller movement and move it along multiple axes at once; (2) Recall the target location and the alignment of the end effector with a fixed structure at the target location; (3) Monitor the camera and window

views to gain a sense of arm location and posture; Or (4) Monitor the arm control GUI on the left-hand monitor to gain a sense of joint angle and the distance to joints limits.

The task simulator (See Fig. 1) consists of six-DOF robotic arm: shoulder pitch, shoulder yaw, elbow pitch, wrist pitch, wrist yaw and wrist roll. It also includes one controller for displacement control and the other is for posture control. Cameras were set to get a comprehensive view of the robotic arm situation. In the simulator there are five cameras: one at the elbow joint, one at the end effector, one at the base and two in the room for global view (See Fig. 2).

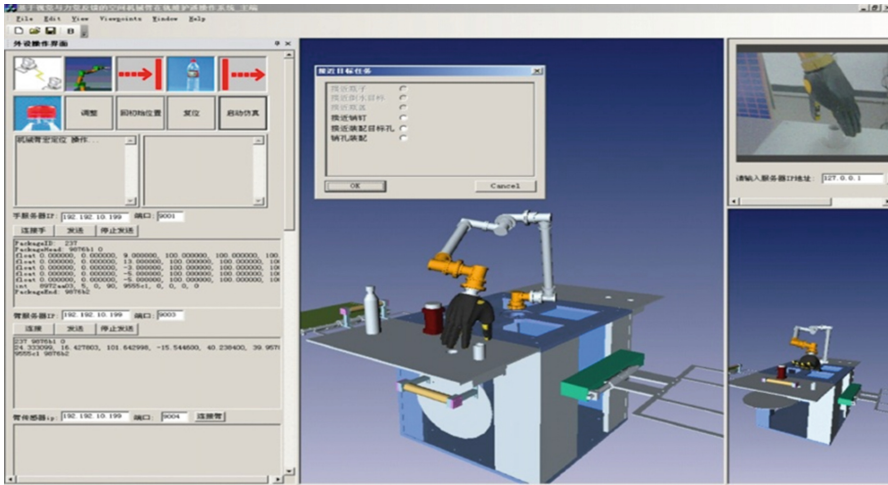


Fig. 1. Information Display and control interface of the teleoperation. *In the middle of interface is a planned routine in virtual reality providing support for control decision and the right above window provides a real view of the robotic arm, which can be changed between the four camera channels.*

2.2 Experimental Design

The independent variable in this study was configuration of cameras and control mode. Each configuration selects each camera from elbow joint, the base, the end effector and the global. In this experiment, seven types of camera configurations were chosen as seven levels of the independent variable. And there are two levels of control modes: manual mode and supervisory mode, by the latter of which tele-operators control the arm under the planned routine. Participants finish the fly-to-catch task based on virtual reality interface. Participants first attended two 2.5 h training sessions on two consecutive days to learn about the concepts and skills required for operating the arm. During the training, the participants performed 3-4 fly-to-tasks with the guidance of allocated routines in virtual reality. All participants were able to complete the training successfully, i.e. they completed the last fly-to-task within tolerance limits (2 meters for translational movement along any axis, 30 degrees for rotations about any axis). On the third day, participants firstly performed three fly-to-catch tasks under each camera

configuration with the virtual reality blocked. On then operators repeat three fly-to-catch tasks under each camera configuration under the guidance of routine allocation shown on virtual reality window. To counteract learning effects, the order of the experimental conditions, and the tasks associated with each condition, were counter-balanced between participants. After the experimental sessions, the participants filled out a debriefing questionnaire to indicate their preferred camera configuration, the strategies they used to perform different steps of the tasks, the main focused view window, and any other comments and suggestions about the designs and experiment.

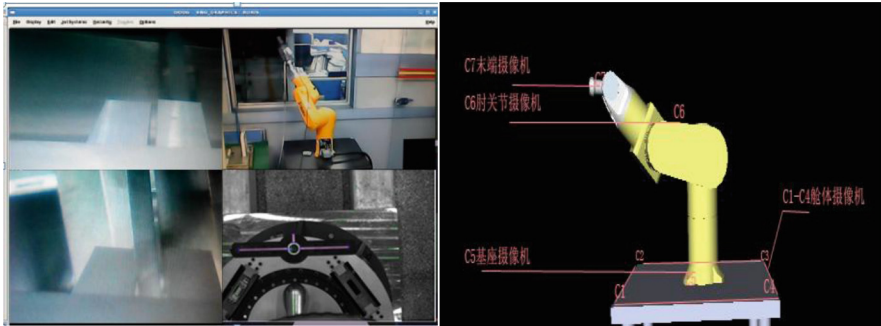


Fig. 2. Task scenario view. Left: views from four selected cameras; Right: camera placements in the remote environment.

3 Results

The data were analyzed with repeated measures ANOVAs using general linear models (SPSS19.0). Overall task time, times of singularity and times of collision warning are collected as the performance results to indicate the efficiency and effectiveness of each tele-operation. The significance of performance difference between supervisory control and manual control is not obvious ($p > 0.05$). As Fig. 3 shows, operators perform best in task time when shown four channels of views from cameras at elbow, base, end-effector as well as the global while the task time is the most when there is no view from the base camera. It seems that there is no too much performance difference when the elbow camera is unused. It is the same case with times of singularity and times of collision warning. However, in term of collision warnings, elbow camera and base camera do not make any difference, especially when left global camera is used although base camera preceded elbow camera when right global camera is applied (Fig. 4).

Eye tracking data from five participants had to be excluded from the analysis due to either low quality of the signal (in one case because the participant was wearing glasses, in another case because of large head movements) or because of calibration issues. The scan patterns of the remaining 7 participants were analyzed in terms of the percentage of dwells (defined as fixations of duration of more than 1 s) on the arm configuration indicators on the Arm Control GUI. Trends were observed suggesting more fixations on views from end-effector and global camera.

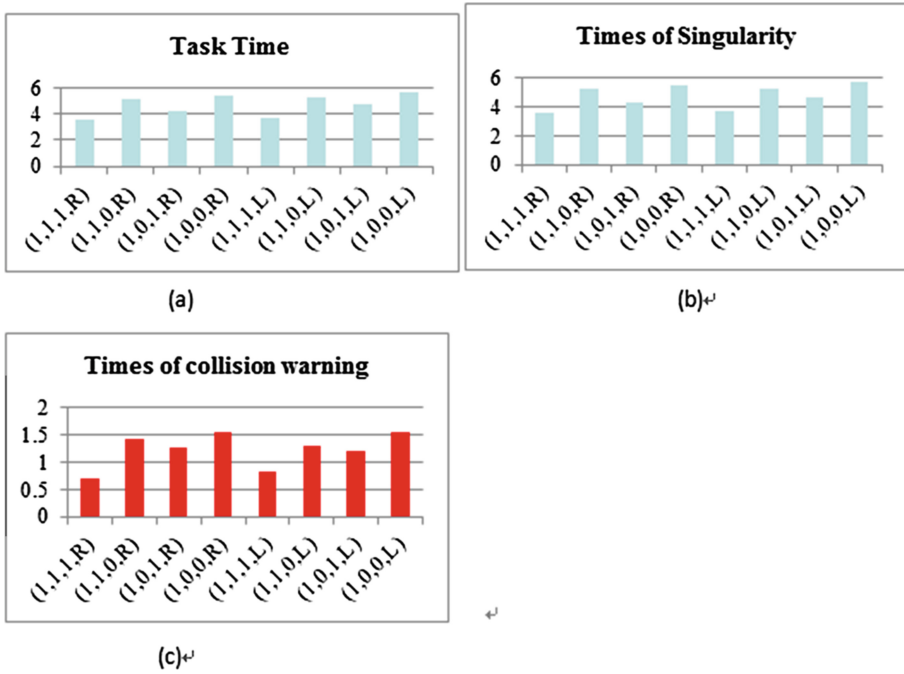


Fig. 3. Teleoperation Performance within seven camera configuration. The four camera positions consist of one configuration set such as (1,1,1,R). The first variable indicates application of the end-effector camera with one level named as “1”; the second variable for elbow camera with two levels—used as “1” and unused as “0”; the third variable for base with two levels—used for “1” and unused as “0”; the fourth variable for global-view camera with two levels—“R” for the one to the right of the arm and “L” for the one to the left of the arm.

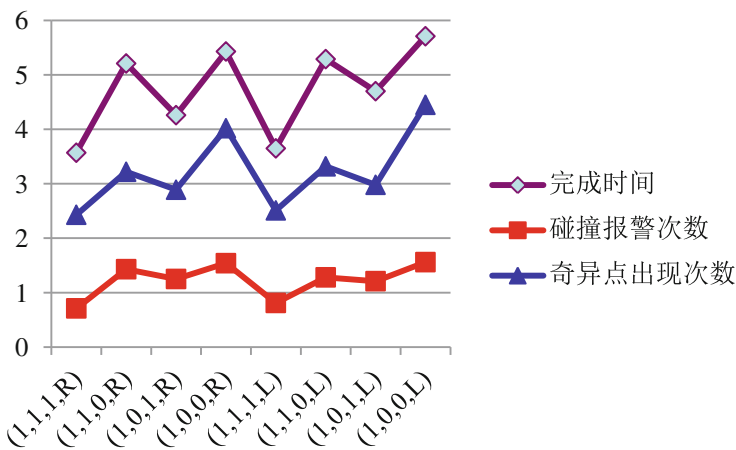


Fig. 4. Comprehensive performance results comparison among seven camera configurations

In the debriefing questionnaire, participants were asked to rank the effectiveness of all the five cameras for supporting operators in monitoring the arm. Ten participants (50 %) preferred the base camera rather than the elbow camera while two cannot make decisions on the preference. Seven participants consider global camera as most essential during fly process while end-effector camera is important at last stage of the task- catch. One participant claimed that elbow camera is useful at the last stage. Five participants claimed that workload is quite high during the fly process when the global camera is blocked and operators depend on base camera or elbow camera to figure out the position and posture in global coordinate.

4 Discussion and Conclusion

Space teleoperation is an especially challenging task due to the complexity and mobility of the system and limited opportunities for direct viewing of the operating environment. Maintaining awareness of the robotic arm's configuration continues to be a major challenge. This study aimed at comparing the effectiveness of seven camera configurations. Interestingly, the operators mainly depend on global camera to sense the position of the robotic arm and the times of singularity increase when there are blocks on the global camera. At this time, the operators normally take great efforts to retrieve the space memory to figure out the position and posture of the robotic arm in the global coordinate. Global camera is especially useful during the fly process while the camera at the end-effector is quite efficient during catch task. Thus global camera and end-effector camera are essential for fly-to-catch task. In choosing between base camera and elbow camera, most operators have bigger preference to base camera especially for efficient performance and to decrease singularity. The elbow camera seems as a supplementary to base camera or global camera. Although four channels of view require more attention allocation and increase the workload of the operators but is still the most efficient way. It is more reasonable that more channels give enough freedom for operators to choose appropriate view angle in the task process. More interestingly, in the experiment task, the supervisory control mode seems not more efficient than the manual mode. One reason may lie in that the task is easy. More researches are expected to explore how the results are applicable to other kinds of tasks.

References

1. Gilliland, K., Schlegel, R.E., Shehab, R.L. Remote manipulator system operator training and performance metrics. In: Proceedings of the Bioastronautics Investigators Workshop, Abstract #031P (2001)
2. Currie, N.J., Peacock, B.: International space station robotic systems operations-a human factors perspective. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. vol. 46(1), pp. 26–30. SAGE Publications (2002)
3. Sheridan, T.B.: Telerobotics. *Automatica* **25**(4), 487–507 (1989)
4. Sheridan, T.B.: Teleoperation, telerobotics and telepresence: a progress report. *Control Eng. Pract.* **3**(2), 205–214 (1995)

5. Vertut, J., Coiffet, P.: Teleoperation and robotics - (Robot technology; v 3A). Cogan Page, London (1985)
6. Fong, T., Thorpe, C., Baur, C.: Multi-robot remote driving with collaborative control[J]. *Industrial Electronics, IEEE Transactions on* **50**(4), 699–704 (2003)
7. Duchon, A.P., Warren, W.H.: A visual equalization strategy on locomotor control of honeybees, robots, and humans. *Psychology Science* **13**, 272–278 (2002)
8. Darken, R., Kempster, K., Peterson, B.: Effects of streaming video quality of service on spatial comprehension in a reconnaissance task. In: *Proceedings of the meeting of I/ITSEC* (2001)
9. Cosenzo, K.A., et al.: Implementation of automation for control of robotic systems. No. ARL-TR-3808. Army Research Lab Aberdeen Proving Ground MD Human Research and Engineering Directorate (2006)
10. Cosenzo, K.A., Parasuraman, R., Novak, A., et al.: Implementation of automation for control of robotic systems[R]. ARMY Research Lab Aberdeen Proving Ground MD Human Research and Engineering Directorate (2006)