Design Interface Mapping for Efficient Free-form Tele-manipulation

Achyuthan Unni Krishnan, Tsung-Chi Lin and Zhi Li[1]

***Abstract*— Motion tracking interfaces are intuitive for free-form teleoperation tasks. However, efficient manipulation control can be difficult with such interfaces because of issues like the interference of unintended motions and the limited precision of human motion control. The limitation in control efficiency reduces the operator's performance and increases their workload and frustration during robot teleoperation. To improve the efficiency, we proposed separating controlled degrees of freedom (DoFs) and adjusting the motion scaling ratio of a motion tracking interface. The motion tracking of handheld controllers from a Virtual Reality system was used for the interface. We separated the translation and rotational control into: 1) two controllers held in the dominant and non-dominant hands and 2) hand pose tracking and trackpad inputs of a controller. We scaled the control mapping ratio based on 1) the environmental constraints and 2) the teleoperator's control speed. We further conducted a user study to investigate the effectiveness of the proposed methods in increasing efficiency. Our results show that the separation of position and orientation control into two controllers and the environment-based scaling methods perform better than their alternatives.**

I. INTRODUCTION

Motion tracking devices, like motion capture systems, RGB+D cameras, wearable and hand-held controllers, enable humans to use their natural motions to control remote robots. Such interfaces are intuitive for *free-form* robot teleoperation, because of how easily humans can (learn to) control motion coordination of complex robot platforms [1]. However, efficiently controlling tele-manipulation using motion tracking interfaces is usually difficult, because of the simultaneous control of many degrees-of-freedom (DOFs). For instance, when the teleoperator needs to adjust the pose of a remote object, the intended control of the object's position may cause unintended control of orientation, or vice versa. The control efficiency of tele-manipulation is also limited by the level of accuracy that human motion control can achieve. Difficulty in efficient tele-manipulation increases the task completion time, operation errors, cognitive and physical workload, and effort to learn interface controls, leading to reduced acceptance of these interfaces by novice users.

This paper investigates two types of approaches to improve the control efficiency of tele-manipulation via motion tracking, without compromising on the intuitiveness. Related work in literature has proposed and validated various case-by-case interface designs that improve control efficiency by separating the controlled degrees-of-freedom [2], introducing motion constraints [3] or using motion scaling methods [4]. However, prior research hasn't systematically evaluated and compared the many variants of these two

[1] Robotics Engineering, Worcester Polytechnic Institute, Worcester, MA 01609, USA {aunnikrishnan, tlin2, zli11}@wpi.edu

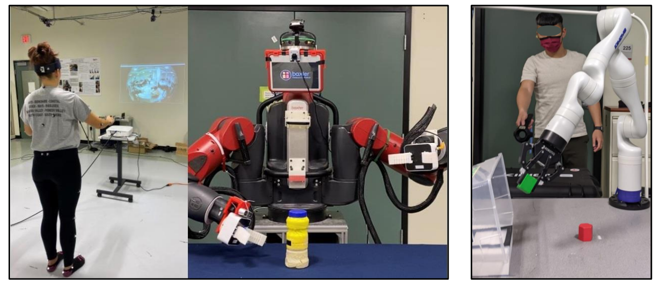

Fig. 1: Robot teleoperation via motion tracking interfaces: Vicon motion capture system (Left); The HTC Vive hand-held controller (Right).

types of approaches, and concluded generalizable interface design theories. To address this limitation, this paper will compare several approaches to separation of controlled DOFs and motion scaling preferred by users in our pilot study, evaluate their performance and workload in the free-form tele-manipulation of 3D objects, and examine the factors that influence the users' preference. The main contribution of this paper is to develop a framework of methods to improve the efficiency of motion tracking tele-manipulation interfaces, and validate this framework in user studies.

II. RELATED WORK

Related work in robot teleoperation and 3D virtual object manipulation have provided many human motion tracking interfaces suitable for intuitive control of free-form tele-manipulation tasks. These interfaces include expensive motion capture systems [1], exoskeletons [5], wearable suits and data gloves [6], [7], and the less expensive but more portable options such as joysticks/gamepads [8], hand-held controllers (of virtual reality systems) [9], RGB+D cameras that capture motions and gestures [10], and touch-screens [11]. These interfaces may capture multi-finger coordination [12], arm-hand coordination [13], bi-manual coordination [1] and whole-body coordination [1], that teleoperates manipulators, mobile, and humanoid robots, or maneuver 3D objects in virtual or augmented reality. Limitations in accuracy of human motion tracking and the need to control multiple DOFs may result in the interference of intended and unintended motions, resulting in reduced operator control efficiency. While being intuitive, motion tracking interfaces are generally limited in their control efficiency, because human motions have limited accuracy, and simultaneous control of many DOFs may cause the interference of intended and unintended motions. Scaling down of speeds and ranges of controlled motion while using motion tracking interfaces can also improve control efficiency by increasing precision as required thus reducing operational errors. The scaling ratios can be fixed (commonly used by tele-robotic laparoscopic or eye surgery interfaces [14]), or vary with the user's operating speed (e.g., PRISM method [15]) or regions of operation [4]. Alternatively, the efficiency of the controlled motions can be improved by introducing constraints in terms of virtual fixtures [16] or autonomy for teleoperation assistance (e.g. collision avoidance [3], motion guidance toward intended goal [1]). From a more general perspective, the interfaces that map gestures or point-and-click actions to autonomous robot motions or movement primitives [2] can all be considered as some kind of constraints that limit the extent to which human operator can control the robot motion freely. In addition, motion constraints can also be introduced by the separation of degrees of freedom (DOFs) in the design of interface mapping. For instance, people may use separate controllers to manipulate a 3D object's position and orientation, to avoid the interference of intended and unintended motion control [2]. Some interface hardware, such as the the trackpad of hand-held controllers, the joystick of gamepads, are naturally suitable for the separation of DOFs as they can clearly distinguish the control inputs for different motion directions. For screen- or projection-based interfaces, interactive avatars such as the ring-and-arrow markers [17], the virtual handlebar [18] enable the independent control of individual DOF(s) of the manipulated (virtual) object or robot end-effector.

Overall, our insight is that the efficiency of tele-manipulation interfaces can be improved by introducing appropriate motion *scaling* and *constraints*. Instead of developing (assistive) autonomy, which has been addressed in abundant recent related work, this paper will investigate two open research questions regarding the **design of interface mapping** for motion tracking interfaces: **RQ1**) How to appropriately separate the controlled DOFs, and **RQ2**) How to adjust the motion scaling ratio. For generalizable results, our proposed interface designs will be implemented on the hand-held controller of a HTC Vive Virtual Reality system, which is a representative motion tracking interface that integrates free-form motion control (via hand motion tracking), directional motion control (via trackpad) and discrete state switching (via buttons). Our evaluation will use a general-purpose task that involves the free-form control of precise reaching, grasping and placing of small objects, which are the skills required in many tele-manipulation task scenarios.

III. Interface Design and Process

This section describes our proposed designs of interface mapping to improve the control efficiency of motion tracking interfaces. We used an interactive design process leveraging findings from the pilot study in cohesion with the implementation of interfaces to continuously improve the design.

A. Primitive Designs of Interface Mapping

Our interactive design process starts with the baseline design and the seven primitive designs of interface mapping that addresses the control efficiency problem of motion tracking interfaces. The **baseline** design is to use a hand-held controller of the HTC Vive Virtual Reality (VR) system to control the 6-DOF pose of the end-effector of the teleoperated manipulator robot. To concisely describe these interface designs, we will use *ND* for non-dominant hand, *D* for dominant hand, *F* for the control using free-form motion tracking, *T* for the control using trackpad (see Fig. 2), *P* for position control; *O* for orientation control. The interfaces used in the final user study in addition to the baseline interface are in bold. The primitive interface designs for the separation of DOFs are:

- **P(D,F) + O(ND,F)**, where 3-DOF position and 3-DOF orientation control of robot end-effector is separated across two controllers in dominant and non-dominant hands respectively (refer Fig. 2);
- **P(D,F) + O(D,T)**, which uses the controller held in the dominant hand to control the 3-DOF position using free-form motion tracking and control the 3-DOF orientation using the trackpad (refer Fig. 2).
- P(D,T) + O(D,F), which uses the controller held in the dominant hand to control the 3-DOF position using the trackpad and control the 3-DOF orientation using free-form motion tracking.
- Mode switch, which uses one hand-held controller held in the dominant hand and allows switching between the three modes of Baseline or P(D,T) or O(D,T);

The primitive interface designs for motion scaling are:

- **Environment-based Scaling** ($Scale_E$) increases the scaling-down ratio of the robot end-effector motion as it approaches environment constraints or targeted objects. $Scale_E$ changes the scaling ratio between the operator's control speed and the controlled robot speed from 1:1 to 5:1 when the robot arm is close enough to (20 cm above; refer Fig. 2) the table surface;
- **Operation-based Scaling** ($Scale_O$) increases the scaling-down ratio of the robot end-effector motion as the operator's control slows down. $Scale_O$ changes the controller-robot scaling ratio from 1:1 to 5:1 when the operator's control speed drops below 0.05 m/s. This method considers the operator's slow controller motion as the intent to perform precise manipulation. After some users couldn't perceive the scaling-ratio of 3:1, the final ratio 5:1 for more noticeable motion scaling for $Scale_O$ and $Scale_E$ was finalized for the user studies;
- Handle-Bar Scaling ($Scale_H$) defines a virtual handlebar by tracking the 3D position of two hand-held controllers, and allows the operator to adjust the scaling ratio by varying the distance between the two hand positions [19]. The $Scale_H$ interface changes the controller to robot scaling ratio from 1:1 to 3:1 when the distance between controllers goes below 10 cm. Thus, the position of robot end-effector can be controlled by the translation of the handlebar between two controllers, while the orientation can be adjusted by rotating the controller held in the dominant hand.

B. Interactive Design Process and Findings

We recruited four pilot study users (3 males and 1 female) to continuously test and improve the primitive interface designs. The participants were asked to teleoperate a 7-DOF manipulator robot (Kinova Gen3) with a two-fingered Robotiq gripper (2F-85) using the hand-held controller of a HTC Vive Virtual Reality (VR) system. As shown in Fig. 2, the tele-manipulation task is to pick up small wood blocks on a remote counter workspace and to place them in a target location inside of a plastic bin. Each of the interfaces was repeated for three trials. The interviews with the pilot study users helped us to identify the effective design features and factors that influence their decision. From the interactive design process, we find out that:

- **Separation of position and orientation control** is generally preferred, yet it is unclear whether the position and orientation control should use the *same or separate controllers*. While some users prefer to control one robot arm with one controller to afford bimanual operation, others prefer separating the position and orientation control to two controllers so that they could perform **simultaneous rotation and translation**. *Trackpad vs hand tracking for orientation control* was also compared in our pilot study, and the user preference was not consistent. Some users preferred hand motion tracking (as in the P(D,F)+O(ND,F) mode) since it felt more natural to rotate the robot end-effector as an object, while others prefer to use trackpad to control orientation so that they can precisely adjust the object rotation in each DOF.
- **Mode Switching** should be avoided because users felt it tedious to switch between different control modes. They may even sacrifice performance to avoid switching modes as they felt it interrupted fluent operation.
- **Motion Scaling**, as we observed in the pilot study, involved the design of scaling ratio and the method to adjust it. When comparing the trackpad vs motion tracking for position control, we found that the pilot users consistently prefer not to use the trackpad for position control as in the P(D,T) + O(D,F) mode, probably because the same scaling ratio doesn't work well for manipulation in both large and small scale. While the trackpad can precisely adjust the end-effector positions for picking and placing, position control of the same scaling would be slow and tedious to control reaching and moving motions across the workspace. Regarding how to adjust the scaling ratio, we found the handlebar ($Scale_H$) [18] was difficult to use because it was hard to maintain the distance between two hand-held controllers required to maintain a steady scaling ratio. The environment-based scaling ($Scale_E$) and the operation-based scaling ($Scale_O$) were both considered useful, but it was unclear which mode would work better in what situation. Moreover, although the autonomous dynamic scaling is more convenient (e.g., $Scale_O$ and $Scale_E$) than manually adjusting the scaling ratio (e.g.,$Scale_H$), it is also important that they keep the scaling ratio to be steady for certain operations (e.g., steady low operator-robot scaling-down ratio for moving towards objects) and consistent in regions of the workspace (e.g., near the table constraints).

The feedback from the pilot study informs our final interface designs to be compared in our controlled user study.

IV. Experiment

We conducted a user study with (N=8) participants to evaluate the usability of the effective interface designs identified in the interactive design process described in the previous section. Specifically, we compared the users' performance and workload in dexterous manipulation using different DOF separation and motion scaling designs. We also examined the preferred combination of interface design features.

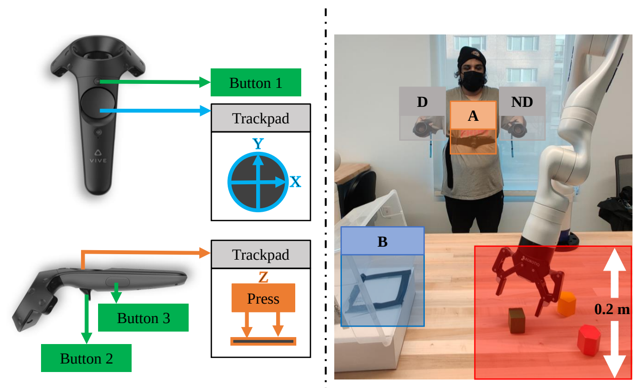

Fig. 2: Control functions (left): 1) Homing robot at the start of teleoperation, or engage/disengage the robot; 2) Toggle to open and close the gripper; 3) Confirming start of robot teleoperation. Touching the trackpad in X- and Y-directions controls the motion along X- and Y-axis; pressing trackpad controls Z-axis motion. The experiment procedure (right) with the region of $Scale_E$ in red, body tracker (A), target location (B), controllers in the dominant (D) and non-dominant (ND) hands.

Participants and Tasks: Our user study recruited (N=8) participants (7 male and 1 female, age = 26.4 $\pm$ 4.2 years). The participants recruited for this final user study are different from the pilot user study. The experimental task is the same task as described in Section III-B and Fig. 2. This task consists of general-purpose manipulation actions across the workspace (e.g., reaching, moving) as well as the precise grasping, and placing of small objects. The plastic bin is placed sideways to force the operator to carefully maneuver the orientation of robot end-effector which requires efficient and precise dexterous manipulation and orientation control.
Experiment Procedure: This experiment consists of six sections. Each section has a *training phase* to familiarize the participants with the interface design selected for that section, and a *task-performing phase* in which the participants officially perform manipulation tasks. The participants practiced and performed the tasks using the baseline interface design in Section 1, and using the four modes of interface designs (described in the Experiment Conditions below) in Sections 2-5. The interface modes for Sections 2-5 were randomized to avoid learning effect. In the last section, the participants combined their preferred mode for DOF separation and motion scaling, and used this combined interface. Total trials for each subject = 6 interfaces $\times$ 3 trials.

Experiment Conditions: Through the interactive design process P(D,F) + O(D,T) and P(D,F) + O(ND,F) for DOF separation, and the $Scale_E$ and the $Scale_O$ interfaces for motion scaling were selected for evaluation.

Data Collection and Analysis: During the experiment, we collected data to analyze the performance and workload to compare the interface design. For task performance, we measured the task completion time, and the occurrence and types of errors. The errors include: 1) collisions with the table and the object, and 2) misplacement of the object outside of the designated region. We also estimated the physical workload from the teleoperator's torso movement tracked by a Vive tracker attached to the operator's chest. Large torso motions indicate that the operator had to move their body to compensate the limited arm motion range, due to the interface design limitations. In order to evaluate the user perceived workload, the NASA-TLX questionnaire was used while a System Usability Scale (SUS) survey was used to evaluate the participants' rating of the interfaces' usability.

The Wilcox rank-sum test was used to identify significant differences between the different NASA-TLX and SUS responses while one-way ANOVA analysis was used to identify significant differences between the interfaces for average task completion times and body motion.

V. Results

In the following sections, the "combination" mode refers to the mode where participants were allowed to use their preferred designs for scaling and separation of DOFs. In the figures presenting the results, we address the interfaces as follows: A = Baseline Design; B = P(D,F) + O(ND,F); C = P(D,F) + O(D,T); D = $Scale_E$; E = $Scale_O$; F = Combination of preferred mode of DOF separation and motion scaling. Additionally, the bars with asterisk on top in Fig. 3 represents a significant difference between the operation modes at the line's starting and ending points.

A. Task Performance

Fig. 3 a) shows the task completion time averaged across the 8 participants for each interface design. We found that based on the one-way ANOVA analysis using an additional controller for orientation control , i.e., the P(D,F) + O(ND,F) mode, significantly reduced the task completion time, compared to the baseline interface ($F(1,46)=11.63$, $p<0.05$), and the interface using trackpad for orientation control, i.e., P(D,F) + O(D,T) ($F(1,46)=5.3$, $p<0.05$). Regardless of what their preference was, the combination mode significantly reduced the task completion time compared to the baseline ($F(1,46)=13.06$, $p<0.05$). Neither $Scale_E$ nor $Scale_O$ significantly reduced the task completion time compared to the baseline mode. This could be because robot motion was slowed down for periods of operation due to scaled motion.

We also analyzed the amount of body motion incurred when using each interface mode. The motion of the body tracker, as attached to the operator in Fig. 2, was used to monitor the operator's torso movements while teleoperating. More body motion during teleoperation implies a worse interface design, because the arm and hand motions for robot control have limited ranges and have to be compensated by (unconscious) movements of the body. As shown in Fig. 3 b), the P(D,F) + O(ND,F) mode, caused significantly less body motions compared to the baseline ($F(1,46)=7.31$, $p<0.05$), and P(D,F) + O(D,T) ($F(1,46)=4.92$, $p<0.05$). Overall, the separation of DOFs were observed to significantly reduce body motion. While scaling down the robot control motions may limit the operator's workspace, we did not find significant difference when compared with the baseline interface. The combination mode caused body motion comparable to the $Scale_E$ and $Scale_O$, which implies motion scaling in general may induce more body motions of the teleoperator.

In Fig. 3 c), we compared the total occurrence of each type of error for all the participants, because the mean and variance of each type of error was very small. The three types of errors we considered include collisions with the table, the objects, and misplacement of the object in the bin. $Scale_O$ caused more collisions with the table compared to all the alternative designs with an average of 1.25 ± 0.33 times per trial, while $Scale_E$ caused 0 collisions with the table across all the participants. $Scale_E$ also caused the least collisions with the object with an average of 1 ± 0.61 times per trial, while the baseline had the most with an average of 2.75 ± 0.78 times per trial. The placement errors for all the interface modes were comparable to the baseline. The combinations of preferred interface designs were neither the most nor the least number of errors, for all types of errors.

B. Survey Feedback

Fig. 3 d) and e) show the results of our NASA-TLX and SUS surveys. We compared the weighted NASA-TLX scores to measure the subjective workload across interface modes. The weighting coefficients selected according to the NASA-TLX guideline questions are: mental demand=5, physical demand=1, temporal demand=0, performance=4, effort=3, frustration=2. We found that none of the 6 interfaces had a significantly different score when compared against the baseline, indicating that all the interfaces have comparable perceived workload. The baseline interface scored the lowest perceived workload when averaged across the 8 participants with a total of 24.25 ± 8.9 while the $Scale_O$ interface scored the highest with a total of 48 ± 12.58. From the SUS survey, we found that the overall usability of the $Scale_O$ is significantly lower ($p<0.05$) than the baseline mode. A more detailed usability analysis indicates that the $Scale_O$ mode is significantly worse ($p< 0.05$) than the baseline mode because of unnecessary complexity.

Fig. 3 f) shows the combination of preferred interface design features for each user, and the extent of their preference. The preferred combinations of interface designs for scaling and DOF separation fell into three types, with a ratio of 3:3:2. Most participants (6 out of 8) preferred environment-based scaling ($Scale_E$) over the operation-based scaling ($Scale_O$), while 5 out of 8 participants preferred the P(D,F) + O(ND,F) interface when it came to DOF separation. No participant

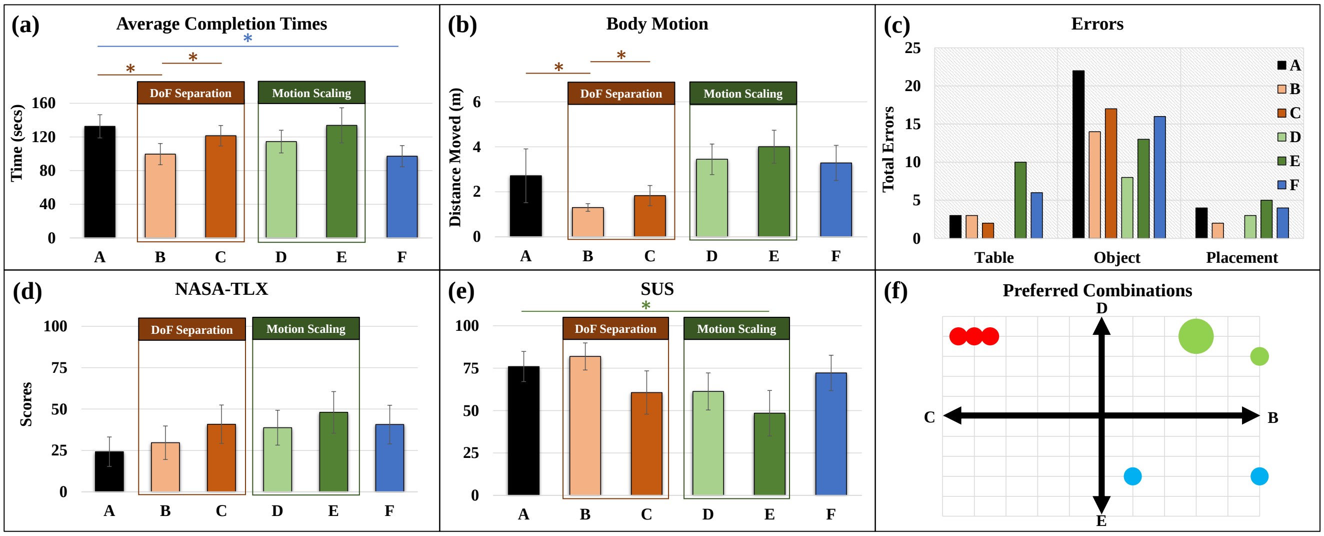

Fig. 3: a) Average task completion times; b) Body motion; c) Task errors; d) Weighted NASA-TLX scores; e) Weighted SUS scores; f) The preferred choices of DOF separation and motion scaling modes. The one larger dot represents two users with same preferences.

preferred the combination of operation-based scaling and the mode using trackpad for orientation control.

VI. Discussion

A. About DOF Separation

Regarding Research Question **RQ1**, the separation of position and orientation control to two controllers is preferred over the alternative method. Overall, DOFs separation can improve control efficiency of a teleoperation interface because it enables the operator to exactly control which DOFs to be involved, and avoid the interference of unintended motions. The separation of position and orientation control to two controllers performs better because it balanced the precision and efficiency in motion control. For the pick-and-place task in our user study, separately controlling the 3-DOF for position or orientation leads to more efficient object manipulation. Using this interface design, the task completion time was significantly reduced by 18% compared to the alternative DOF separation method and by 25% compared to the baseline interface. Ability to control individual DOF may be necessary for the efficient control of translation (e.g., object inserting, stacking), or efficient adjustment of orientation (e.g., adjusting the pan or tilt of an active telepresence camera). Particularly, the body motions observed in the usage of the two-controller interface, P(D,F) + O(ND,F), is 29% less compared to mode using trackpad for orientation control, and 58% less compared to the baseline. Note that induced body motion is not preferred as it increases a teleoperator's physical fatigue and frustration. Since the operator does not have to control both the controllers simultaneously to teleoperate the robot, separation of rotation and orientation across two controllers does not increase the cognitive workload. This is highlighted by the survey feedback, where separating the position and orientation motion control did not significantly increase the perceived mental workload, and had the best usability score. Since using trackpad for orientation control is not as intuitive, 5 out of 8 participants preferred the two-controller interface over the alternative for DOF separation.

B. About Motion Scaling

Regarding the research question **RQ2**, we found that the environment-based scaling performed better and was preferred for our user study task. This is because the environment-based scaling improves the precision of motion control about environment constraints, or the local region where the precise manipulation is performed. This results in reduced errors and unintended motions, thus increasing interface efficiency. In general, environment-based motion scaling may also benefit dexterous manipulation or navigation in a cluttered environment, similar to methods for collision avoidance. As shown in Fig. 3 c), the environment-based scaling, i.e., the $Scale_E$, caused fewer errors than the operation-based scaling method ($Scale_O$). The environment-based scaling is particularly effective in reducing collisions with objects when the participant attempted or failed to grasp. It also reduced the task completion time by 14% and 13% compared to $Scale_O$ and the baseline, respectively. We noticed that compared to baseline interface, $Scale_E$ has comparable perceived workload and usability score, while those of the $Scale_O$ are significantly worse. Our survey shows that 6 out of 8 participants preferred the environment-based scaling over the alternative method. The post-study interview shows that when using environment-based scaling, participants could better predict when the scaling ratio will change. However, the operation-based scaling, which is much less predictable and controllable, usually lead to unintended changes of scaling ratio.

C. Combining Two Types of Design

We also found that the combination of a user's preferred choices for DOF separation and motion scaling may lead to a balanced performance and workload. The combination of preferred methods, regardless of individual difference in the choices has comparable task completion time compared to the baseline, even if for some part of the manipulation operation the controlled motion became slow due to the motion scaling. The survey feedback shows that the usability score of the combined method is comparable to the baseline interface. For the task in this user study, a balanced performance and acceptable workload may contribute to improved user's acceptance, yet the choices of preference may be different for individuals and other tasks and types of manipulation.

VII. Conclusion and Future Work

This paper proposed methods for the separation of Degrees of Freedom (DOF) and motion scaling and evaluate whether they can effectively improve control efficiency in tele-manipulation using motion tracking interfaces. The user study results and survey feedback show that it is preferred to separate the position and orientation control to two separate controllers, and to adjust the motion scaling based on the regions of operation instead of the operation speed. Essentially, the preferred methods for DOF separation avoids the interference of unintended motions, while the preferred motion scaling methods improves control precision about the environment constraints and in the regions of interest, reducing errors and wasted motion. Although the proposed methods increased the complexity of the teleoperation interface to some extent, they did significantly improve the task performance without introducing significantly more mental workload. Our work is limited because of the small sample size and simple task design. Verifying the results of this paper with a larger population is part of our intended future work. There are many other dexterous manipulation tasks to be considered, which may require efficient manipulation interfaces like inserting, stacking, boundary-tracing, and precise orientation control. Since the primary application of the proposed technology is for nursing robot teleoperation, our future user studies will include tasks in the context of nursing assistance and a larger number of participants from population of registered nurses and nursing students. It is possible that the performance and workload of the proposed methods may be different for other types of manipulation tasks and user population.

References

[1] T.-C. Lin, A. U. Krishnan, and Z. Li, "Intuitive, efficient and ergonomic tele-nursing robot interfaces: Design evaluation and evolution," *Accepted by the Transactions on Human-Robot Interaction*, 2021.

[2] T. Zhou, M. E. Cabrera, J. P. Wachs, T. Low, and C. Sundaram, "A comparative study for telerobotic surgery using free hand gestures," *Journal of Human-Robot Interaction*, vol. 5, no. 2, pp. 1–28, 2016.

[3] N. Pedemonte, F. Abi-Farraj, and P. R. Giordano, "Visual-based shared control for remote telemanipulation with integral haptic feedback," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 5342–5349.

[4] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge, "A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments," in *Computer graphics forum*, vol. 38, no. 1. Wiley Online Library, 2019, pp. 21–45.

[5] F. Falck, K. Larppichet, and P. Kormushev, "De vito: A dual-arm, high degree-of-freedom, lightweight, inexpensive, passive upper-limb exoskeleton for robot teleoperation," in *Annual Conference Towards Autonomous Robotic Systems*. Springer, 2019, pp. 78–89.

[6] G. A. Yashin, D. Trinitatova, R. T. Agishev, R. Ibrahimov, and D. Tsetserukou, "Aerovr: Virtual reality-based teleoperation with tactile feedback for aerial manipulation," in *2019 19th International Conference on Advanced Robotics (ICAR)*. IEEE, 2019, pp. 767–772.

[7] S.-K. Kim, S. Hong, and D. Kim, "A walking motion imitation framework of a humanoid robot by human walking recognition from imu motion data," in *2009 9th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2009, pp. 343–348.

[8] Z. Li, P. Moran, Q. Dong, R. J. Shaw, and K. Hauser, "Development of a tele-nursing mobile manipulator for remote care-giving in quarantine areas," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 3581–3586.

[9] D. Rakita, B. Mutlu, and M. Gleicher, "A motion retargeting method for effective mimicry-based teleoperation of robot arms," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 361–370.

[10] I. Ajili, M. Mallem, and J.-Y. Didier, "Gesture recognition for humanoid robot teleoperation," in *2017 26Th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 2017, pp. 1115–1120.

[11] Y. Cho and F. L. Hammond, "Improving efficiency and safety in teleoperated robotic manipulators using motion scaling and force feedback," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2020, pp. 1236–1242.

[12] A. Handa, K. Van Wyk, W. Yang, J. Liang, Y.-W. Chao, Q. Wan, S. Birchfield, N. Ratliff, and D. Fox, "Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 9164–9170.

[13] S. Li, J. Jiang, P. Ruppel, H. Liang, X. Ma, N. Hendrich, F. Sun, and J. Zhang, "A mobile robot hand-arm teleoperation system by vision and imu," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 10 900–10 906.

[14] J. Guo, C. Liu, and P. Poignet, "A scaled bilateral teleoperation system for robotic-assisted surgery with time delay," *Journal of Intelligent & Robotic Systems*, vol. 95, no. 1, pp. 165–192, 2019.

[15] S. Frees, G. D. Kessler, and E. Kay, "Prism interaction for enhancing control in immersive virtual environments," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 14, no. 1, pp. 2–es, 2007.

[16] C. P. Quintero, M. Dehghan, O. Ramirez, M. H. Ang, and M. Jagersand, "Flexible virtual fixture interface for path specification in telemanipulation," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 5363–5368.

[17] W. Pryor, B. P. Vagvolgyi, A. Deguet, S. Leonard, L. L. Whitcomb, and P. Kazanzides, "Interactive planning and supervised execution for high-risk, high-latency teleoperation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 1857–1864.

[18] P. C. Gloumeau, W. Stuerzlinger, and J. Han, "Pinnpivot: Object manipulation using pins in immersive virtual environments," *IEEE transactions on visualization and computer graphics*, vol. 27, no. 4, pp. 2488–2494, 2020.

[19] P. Song, W. B. Goh, W. Hutama, C.-W. Fu, and X. Liu, "A handle bar metaphor for virtual object manipulation with mid-air interaction," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2012, pp. 1297–1306.