

## RESEARCH ARTICLE

# A Shared Control Framework for Enhanced Grasping Performance in Teleoperation

YAONAN ZHU<sup>ID</sup>, (Member, IEEE), BINGHENG JIANG<sup>ID</sup>, QIBIN CHEN<sup>ID</sup>,  
TADAYOSHI AOYAMA<sup>ID</sup>, (Member, IEEE), AND YASUHISA HASEGAWA<sup>ID</sup>, (Member, IEEE)

Department of Micro-Nano Mechanical Science and Engineering, Nagoya University, Nagoya 464-8603, Japan

Corresponding author: Yaonan Zhu (zhu@robo.mein.nagoya-u.ac.jp)

This work was supported in part by JST Trilateral AI Research, Japan, under Grant JPMJCR20G8; and in part by JSPS KAKENHI under Grant JP22K14222.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of the Graduate School of Engineering, Nagoya University under Approval No. 22-3.

**ABSTRACT** Remote teleoperation has shown significant advancements since the first teleoperation system was proposed by Goertz in the 1940s. In recent years, the research on shared control methodologies in which the robot assists the operators in accomplishing the desired tasks has gained extensive attention. One such important task in teleoperation is object grasping. In this paper, we propose a shared control framework to enhance the teleoperated grasping performance. The proposed framework is built upon a virtual reality device-based direct teleoperation system. In this framework, a template matching-based object point cloud compensation is introduced for multi-angle grasping pose generation. Then, the feasible grasping candidates are selected considering joint constraints-aware manipulability. Finally, the grasping assistance is achieved by trajectory blending with dynamic authority adjustment. To validate the performance of the proposed framework, we carried out experimental evaluations. The output results indicate improved grasping performance in terms of reduced task completion time, linear trajectory, and workload.

**INDEX TERMS** Grasping assistance, point cloud compensation, shared control, teleoperation.

## I. INTRODUCTION

Robotic teleoperation has been widely and extensively studied throughout robotics history since the idea was first proposed by Goertz in the 1940s [1]. Generally, direct teleoperation systems are designed to constantly follow operators' control, and operators receive visual, haptic, or proprioceptive sensory feedback to achieve a remote presence. The quality of sensory perception was later denoted as transparency. The quality of transparency directly affects task performances and control intuitiveness in a direct control framework [2]. However, due to the degraded quality of long-range signal transmission and human-robot embodiment heterogeneity, perfect transparency only exists in ideal situations [3]. This facilitated the development of new approaches which integrate direct teleoperation with

some level of automation on the follower side for operational assistance. The approach is referred to as shared-control [4], in which the human cognitive skills and the robustness of the follower robots are both leveraged [5].

The shared control research has been widely conducted in application areas such as space exploration [6], surgical robotics [7], hazardous material handling [8], multiple robots teleoperation [9], and assistive robotics [10]. In most of the above-mentioned application cases, human operators have to teleoperate complex multi-degrees-of-freedom (DOF) robotic systems to accomplish object manipulations. In this context, object grasping takes on a particularly significant and fundamental role, which serves as a critical component in the success of these operations [11]. However, controlling all DOFs of complex robotic systems to precisely grasp a desired object presents a significant challenge. This difficulty is due to the degraded sensory perception of depth information [12], the requirement for simultaneous control of

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Liu<sup>ID</sup>.

the end-effector's position and orientation, and the presence of robot constraints [13]. To overcome these challenges and successfully execute a grasp, operators have to persistently monitor the condition of the robotic system, which can be both physically and mentally demanding, leading to operators' fatigue and degraded task performance [14].

Hence, it is of great necessity to develop grasping assistance systems to reduce workloads and enhance remote grasping performance. In [15], Abi Farraj et al. introduced a novel haptic shared-control approach, in which a point cloud-based autonomous grasping pose detection algorithm [16] was integrated with the haptic guidance for assisting a human operator in the sorting and segregation of different objects in a cluttered environment. Ghalamzan et al. [17] extended grasping assistance to the post-grasping phase and proposed a shared control approach to help operators select a stable grasping pose by considering post-grasping manipulability. For the non-haptic guidance approach, grasping assistance is also applied to the control of a multi-DOF robotic arm for the target grasping task through a simple non-invasive Brain-Computer Interface and computer vision guidance [18]. Xu et al. designed a system in which the operators only need to send two kinds of signal instructions to achieve the transnational motion of the robot arm, and once the end effector reaches the predefined visual guidance area, the control is switched to full autonomy to complete the grasping. Unlike the approach of switching on-off of automation in the grasping assistance process, in [19], Laghi et al. has developed a grasping assistance algorithm for a bimanual teleoperation system in which the end effector trajectories are blended based on the user's willingness to grasp. The system utilizes flexible Virtual Reality (VR) controllers as input devices and assists the operation through the automatic coordination of bimanual motions to grasp a single object with different sizes. Moreover, rather than assisting in full DOF, Bowman et al. [20] adopted a DOF-wise control authority allocation between the human and the robot to achieve flexible grasping assistance.

Above mentioned grasping assistance algorithms are dependent on the grasping poses generated directly on the perceived point cloud. Suitable grasping poses are detected based on the input point cloud and the internal classification network [16]. However, this makes the quality of generated grasping poses directly affected by the input point cloud. In most cases, the depth camera mounted on the robot can only obtain the point cloud from a single fixed direction either from the side or from the top. For example, in the works of [15], [18], and [19], the depth camera was mounted on top of the workspace to simplify the grasping pose generation and achieve the grasping from above for higher manipulability. However, for cutting-edge multi-purpose robotic platforms such as Tiago (Pal Robotics) [21], Human Support Robot (TOYOTA, HSR) [22], and PR2 (Willow Garage) [23], the depth camera is mounted on the robot head to provide point clouds from the front side of the object. Due to the object self-collision, the grasping poses are generated only in the part of the object that is visible to the depth camera and

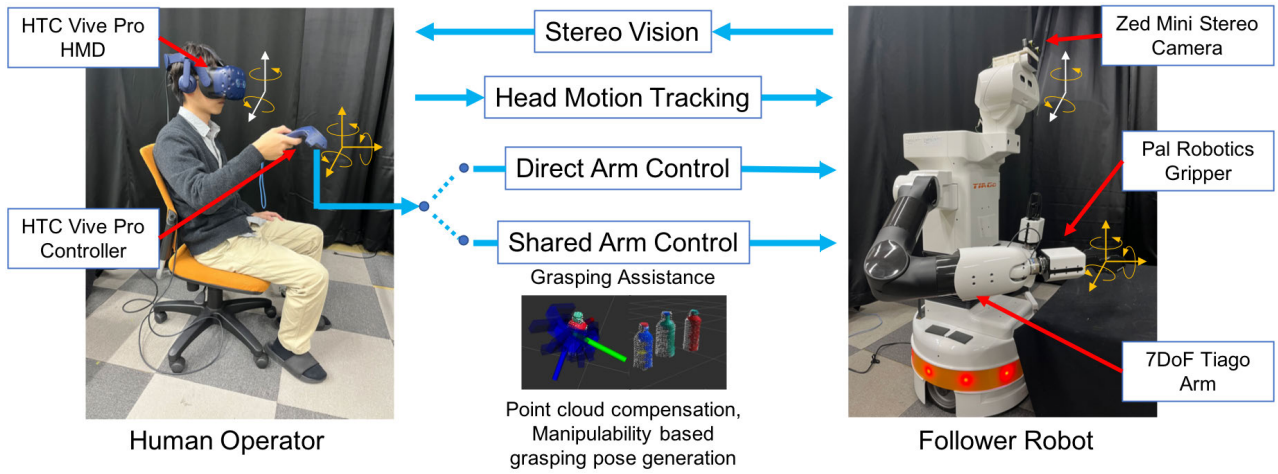
all facing the same direction. This limits the diversity of generated grasping poses and makes it difficult to select suitable assistive grasping poses from multiple angles, which potentially can be used to meet the robot constraints such as joint limits or manipulability [24].

Moreover, a strategy to select an adequate assistive grasping pose from multiple candidates is still an open problem. In human-robot teleoperation systems, due to the embodiment difference between humans and robots, grasping commands that are given by human operators may violate the kinematic constraints of follower robots and make them fall into singularities [21]. Although, most of the previously mentioned grasping assistance systems assumes that the robot approaches the target with high manipulability, in practice, it is important to consider robot constraints and manipulability [24] to guarantee robot execution when generating feasible grasping poses. Thus, to address the challenges of generating multi-directional grasping poses, and avoiding robot constraints for smooth task execution, this paper develops a shared control framework that can generate grasping poses from multiple angles and considers robot manipulability constraints when selecting feasible assistive grasping poses.

On the other hand, compared with conventional fixed-base teleoperation interfaces, VR devices have the advantage of lower cost, flexibility, and extended workspace, which makes it ideal to be utilized as a teleoperation interface [25]. In addition, unlike Liquid Cristal Display (LCD) screens which are commonly used as the visual feedback interface to provide a monocular view [26], VR devices can be integrated with stereo cameras to present a stereo vision to operators through Head Mounted Display (HMD) for improved spatial perception [27]. In this study, a VR device-based teleoperation system is developed. The follower robot is controlled to follow human arm motion intuitively, and a stereo vision is provided for visual feedback with enhanced spatial perception. The proposed shared control framework that assists grasping is built upon this VR device-based teleoperation system.

In this paper, we propose a grasping assistance shared control framework for enhanced grasping performance in a VR device-based teleoperation system. The contributions of this paper are summarised as follows:

- 1) A VR device (HTC, Vive Pro) based teleoperation system is proposed for intuitive direct control. Human arm and head motions are captured by the optical trackers on the Head Mounted Display (HMD) and the VR controller. Captured motions are then mapped to an anthropomorphic robotic manipulator (Pal Robotics, Tiago++) through inverse kinematics with null space resolution. A stereo video stream is displayed on the HMD for stereo visual feedback.
- 2) A shared control framework is proposed for grasping assistance. In this framework, a multi-angle grasping poses generation is achieved by template-matching based point cloud compensation. These multi-angle grasping poses are then used to select feasible grasping



**FIGURE 1.** The proposed shared control framework and the VR device-based teleoperation system. The Head Mount Display (HMD) and the controller of HTC Vive Pro are used as the leader devices to track the operator’s head and hand motion. The 7-DoF robot arm and the parallel gripper of Tiago are used as the follower device. A stereo camera (Stereo Labs, ZED Mini) is used for stereo visual feedback. The shared control method compensates for the objects’ missing point cloud and generates feasible grasping poses based on manipulability. Then a trajectory blending is performed for grasping assistance.

poses by referencing joint limit-aware manipulability measurement. A trajectory blending method with dynamic authority adjustment is introduced to achieve smooth grasping assistance.

- 3) The contribution of point cloud compensation on the improved manipulability distribution over the workspace is analyzed. Moreover, human subject experiments are conducted to evaluate the system’s performance through the metrics of task completion time, linear trajectories, and NASA TLX workload ratings.

## II. VR DEVICE-BASED TELEOPERATION SYSTEM

This section describes the VR device-based teleoperation system that is proposed in this paper. The entire software architecture is built upon Robot Operating System (ROS). To achieve intuitive teleoperation, coordinated control of the robot arm and end-effector according to the operator’s arm and hand motions is important. Hence, as the leader device, a VR device (HTC, Vive Pro) is used to capture human hand and head motions. For the follower device, a redundant robot manipulator with a 7-DOF arm and a 2-DOF head (Pal Robotics, Tiago++) is used to execute the human command. A parallel gripper is attached to the wrist of the robot arm as an end-effector and can be controlled by the trigger of the VR controller. The stereo vision is provided by displaying images captured by the stereo camera (Stereo Labs, ZED mini) on the HMD. Fig. 1 shows the architecture of the proposed teleoperation system.

### A. MOTION MAPPING

In the proposed system, only arm motion mapping is considered. Tiago’s arm has 7-DOFs which is a redundant manipulator that can be used to intrinsically track human arm motions in cartesian space. Tiago’s head has two motors

that can be used to generate pitch and yaw motions to track human head movement. Arm motion tracking and head motion tracking are performed individually.

The end-effector pose in the base frame of the follower robot can be defined as a homogeneous transformation matrix in  $SE(3)$  as follows:

$$T_e^b = \begin{bmatrix} R_e^b & p_e^b \\ 0 & 1 \end{bmatrix} \quad (1)$$

where  $R_e^b$  is a rotation matrix in the robot base frame which satisfies  $R_e^b \in SO(3)$ , and  $p_e^b \in \mathbb{R}^3$  is a translational vector that contains translational information in each cartesian coordinate axis. Similarly, the captured human hand pose is used as the desired end-effector position  $T_{e,d}^b$ .

Firstly hand motion tracking is performed. Since the HTC Vive base station and robot base have different origins in each coordinate system, direct mapping would not be possible. Hence, for translational motion tracking, the relative position with respect to the initial tracker position is used.

$$p_{e,d,k}^b = p_{e,0}^b + (p_{hand,k}^{htc} - p_{hand,0}^{htc}) \quad (2)$$

where,  $k$  represents  $k$ ’th sampling step, and 0 means the onset of the motion. Then, for the rotational motion mapping, the absolute orientation originating from the robot base frame is used as follows:

$$R_{e,d,k}^b = R_{htc}^b \cdot R_{hand,k}^{htc} \quad (3)$$

where a rotation matrix  $R_{htc}^b$  maps the rotation in the HTC coordinate system to the robot base coordinate system.

### B. INVERSE KINEMATICS

Once the desired end-effector pose is obtained from the previous section, an inverse kinematics solver is performed to solve the desired joint angles, which is denoted as:

$$\theta = [\theta_1, \dots, \theta_7]^T \in \mathbb{R}^7 \quad (4)$$

Firstly, the end-effector position and orientation error can be defined as:

$$\mathbf{e}_e = \begin{bmatrix} \mathbf{e}_{e,p} \\ \mathbf{e}_{e,o} \end{bmatrix} = \begin{bmatrix} \mathbf{p}_{e,d} - \mathbf{p}_{e,k} \\ \log(\Delta \mathbf{R}_e) \end{bmatrix} \quad (5)$$

where,  $\log(\Delta \mathbf{R}_e)$  is the logarithm mapping of the orientation error  $\Delta \mathbf{R}_e$  which can be described as  $\mathbf{R}_{e,d} \mathbf{R}_{e,k}^T$ . Note that in the practical implementation, unit quaternions are used to describe the orientation. Here, the task space error  $\mathbf{e}_e \in \mathbb{R}^6$  and the joint space configuration  $\boldsymbol{\theta} \in \mathbb{R}^7$ , which makes the robot manipulator redundant.

Then, in order to track the desired end-effector pose  $\mathbf{T}_e^b$  with Tiago's redundant manipulator arm, an inverse kinematics solver with the null-space resolution is implemented. The general form of the inverse kinematics that contains null space projection can be described as follows [28]:

$$\dot{\boldsymbol{\theta}} = \mathbf{J}^\# \mathbf{e}_e + (\mathbf{I} - \mathbf{J}^\# \mathbf{J}) \boldsymbol{\phi} \quad (6)$$

where  $\dot{\boldsymbol{\theta}} \in \mathbb{R}^7$  is the desired joint velocities, and  $\mathbf{J}^\#$  is the Moore-Penrose pseudo-inverse of the task Jacobian. The Moore-Penrose pseudo-inverse is solved by using the Singular Value Decomposition (SVD) method. The null space projection  $(\mathbf{I} - \mathbf{J}^\# \mathbf{J})$  projects the subsequent vector to the task Jacobian  $\mathbf{J}$ , and the  $\boldsymbol{\phi}$  is a vector that contains the error of the task with secondary priority. In which the vector  $\boldsymbol{\phi}$  can also be interpreted as the desired null space velocity that modifies joint space behavior and does not interfere with the execution of the prioritized task.

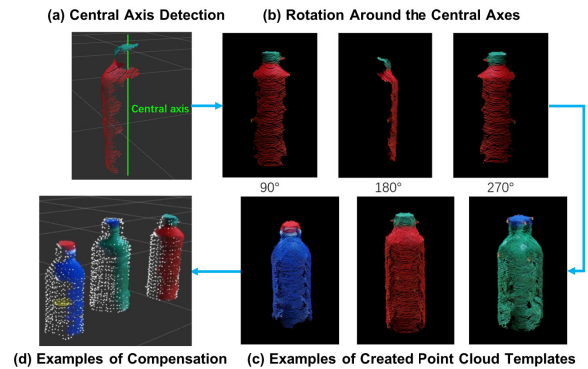
For a redundant manipulator, the null space vector  $\boldsymbol{\phi}$  can be used to optimize the null space behavior based on some criteria such as manipulability indices, joint limits, or joint velocities [29]. In this paper, we employ the null space optimization criterion proposed in [30] and [31] with a preferred arm posture to avoid unreachable task poses. The cost function of the criterion is defined as follows:

$$G(\boldsymbol{\theta}) = \frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\theta}_{prefer})^T \mathbf{K}_w (\boldsymbol{\theta} - \boldsymbol{\theta}_{prefer}) \quad (7)$$

where the  $\mathbf{K}_w$  is a diagonal weighting matrix and  $\boldsymbol{\theta}_{prefer}$  is the preferred arm posture described in joint space. This criterion is computationally simple and has been used to create human-like motions in anthropomorphic robot arms [30]. The null space velocity  $\boldsymbol{\phi}$  can be given by taking the gradients of the cost function (7) in the descending direction. Moreover, the joint limit is implemented by setting a software constraint. This inverse kinematics is implemented under ROS framework with Pinocchio motion library [32].

### III. SHARED CONTROL FRAMEWORK FOR GRASPING ASSISTANCE

In this section, the proposed shared control framework is introduced. The framework is composed of a point cloud compensation-based grasping pose generator, a manipulability-based grasping candidate selection, and a shared control framework with dynamic authority adjustment that complements human commands for grasping assistance. As stated earlier, most of the point cloud-based grasping



**FIGURE 2.** Process of object point cloud compensation. (a) shows the central axis detection. In (b) based on the central axis, the object point cloud is rotated 90°, 180°, and 270°. (c) shows examples of the created templates. (d) shows examples of the compensated point clouds through template matching.

assistance shared control generates grasping pose from only one direction [15], [18], [19]. Hence, we propose the multi-angle grasping pose generation method, in which the diverse grasping poses can be used as candidates to select feasible grasping poses according to manipulability measurement.

#### A. MULTI-DIRECTIONAL GRASPING POSE GENERATION

In this paper, a state-of-the-art grasping pose detection (GPD) algorithm [16] is adopted to generate 6-DOF grasping candidates. However, given a partially observed point cloud of an object, it is unable to calculate the curvature of the missing surface, and some grasp candidates from the side direction are not considered as successful grasp. This constrains the generated grasping poses to only one direction. Hence, the most straightforward way to generate multi-directional grasping poses is to compensate for the object's missing point cloud and then apply the GPD algorithm.

To compensate for the objects' point cloud, an object point cloud library-based template matching is applied. The first step for point cloud compensation is to establish a template library. Since the performance of the template matching can be affected by point cloud distribution characteristics, the depth camera (ASUS, Xtion) is used to create the template library and to do the real-time point cloud compensation.

Given the point cloud of the objects with a symmetric shape, the process of point cloud compensation is shown in Fig.2 and follows the following steps:

- 1) Central axis calculation: The point cloud of the object's top part (cap) is fully obtainable. By segmenting the top part and calculating its central axis, the central axis of the object can be obtained in vector form, as shown in Fig.2 (a).
- 2) Obtain the copies of the point cloud at four different angles: The rotation around the central axis can be applied to the point cloud to approximate the process of obtaining the point cloud from four different angles (Fig.2 (b)).

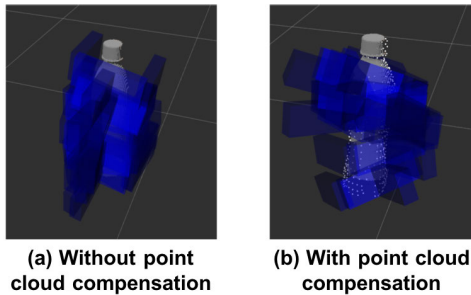


**Algorithm 1** Point Cloud Compensation

---

**Input:** Point Cloud  $P$ , Template Library  $L$   
 Initialization: Output Point Cloud  $O$   
 Segment  $P$  To Get Object-Based Sub-Set  $P^*$   
**for** Object Point Cloud  $p_i$  in  $P^*$  **do**  
 Initialization: Best Matched Template  $M_i$ ,  
 Best Score  $S_i$ , Transform  $T_i$   
**for** Template  $l_j$  in  $L$  **do**  
 Initialization: Matching Score  $S_j$ , Transform  $T_j$   
 $S_j, T_j = \text{MATCH}(p_i, l_j)$   
**if**  $S_j > S_i$  **then**  
 $S_i = S_j, T_i = T_j, M_i = l_j$   
**end if**  
**end for**  
 Apply Transform  $T_i$  to  $M_i$   
 Add The Transformation Result to  $O$   
**end for**  
**Output:**  $O$

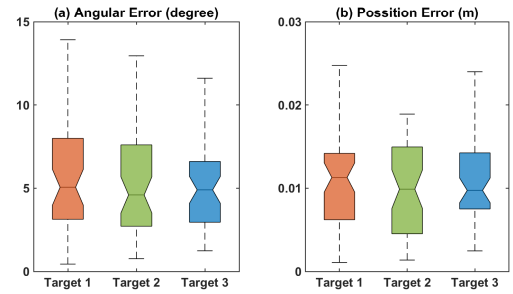
---



**FIGURE 3.** Differences in the distribution of generated grasping poses with and without point cloud compensation. (a) without compensation, the grasping poses are facing the forward direction. (b) with compensation, multi-angle grasping poses are generated.

- 3) Create a template: These rotated point clouds are merged, and voxel grid filtering is applied to remove the overlap before it is finally used as a template of the object. (Fig.2 (c)).
- 4) Point cloud compensation by template matching: A template matching algorithm (Algorithm 1) is performed to compensate for the missing point clouds. The result of point cloud compensation is illustrated in Fig.2 (d).

Algorithm 1 describes the process of point cloud matching and compensation. The input is the real-time point cloud  $P$  obtained by the depth camera, and the object point cloud template library  $L$ . Firstly, the  $P$  is segmented into sub-sets  $P^*$  that contain the point cloud of each object in the scene. Then, the template matching is performed to each point cloud inside the sub-sets  $P^*$ . Here, each segmented point cloud  $p_i$  is used as a template to match each point cloud template  $l_j$  in the library. Based on the highest matching score  $S_i$ , the corresponding transformation matrix  $T_i$  is applied to the matched template  $M_i$  and added to the output point cloud set  $O$ . The Algorithm is implemented under the framework of Point Cloud Library (PCL) [33],



**FIGURE 4.** Error of the object's central axis in the compensated point cloud. (a) the angular error between the compensated central axis and the vertical axis, and (b) the position error between the compensated centroid and the ground truth centroid. Targets indicate three objects shown in Figure 9 (b).

with RANSAC-based SAC-IA function [34] to estimate the transformation matrix and matching scores. The output point cloud set  $O$  is then fed separately into the GPD for grasping pose generation. The difference between with and without point cloud compensation and corresponding grasping poses are shown in Fig. 3. The point cloud compensation for one object can be run within 1s.

Figure 4 shows the error of the object's central axis after point cloud compensation. Target 1, Target 2, and Target 3 represent three different target objects used in the experiments (Figure 9 (b)). Figure 4 (a) presents the angular error between the object's compensated central axis and an axis that is vertical to the table. Since the object is in a symmetrical shape, the angular error can be given by using the axis-angle representation, without considering the rotation around the Z-axis. The median of the angular error is 5.057 degrees, 4.597 degrees, and 4.900 degrees for Target 1, Target 2, and Target 3, respectively. The positional error is shown in Figure 4 (b), which is the Euclidean distance from the object's centroid position in the compensated point cloud to its ground truth centroid position. The median of the positional error is 0.011 m, 0.009 m, and 0.009 m for Target 1, Target 2, and Target 3, respectively.

## B. MANIPULABILITY-BASED GRASPING POSE SELECTION

Among the generated multi-directional grasping poses, some of them may cause the Tiago arm to fall into the singularity, making both prior and subsequent arm motions unfeasible to execute. Therefore, in this section, a manipulability index-based grasping pose selection method is proposed to choose the most feasible grasping candidate.

### 1) MANIPULABILITY CALCULATION

Manipulability index [35] which describes the distance to the singular configuration, is a well-known criterion for determining the ability of robot manipulators to maneuver in the workspace. A larger manipulability value indicates that the robot arm can move smoothly around the corresponding joint configuration. By applying the manipulability index to grasping pose selection, the unfeasible grasping poses can be effectively filtered out. The manipulability measurement is

defined as follows:

$$M(\theta) = \sqrt{\det(\mathbf{J}(\theta)\mathbf{J}(\theta)^T)} \quad (8)$$

where,  $M(\theta)$  is the manipulability value in the joint configuration  $\theta$  corresponding to a grasping pose, and  $\mathbf{J}(\theta)$  is the jacobian matrix under the joint space configuration  $\theta$ .

## 2) MANIPULABILITY WITH JOINT LIMIT PENALIZATION

The equation (8) calculates the manipulability without considering robot joint limits, which can also bring the robot arm to unfeasible configurations. To further ensure the feasibility of the selected grasping poses, this paper follows the suggestion in [36] to introduce the following penalization term that considers the influence of the lower ( $l_j^-$ ) and the upper ( $l_j^+$ ) joint limits:

$$P(\theta) = 1 - \exp\left(-k \prod_{j=1}^n \frac{(\theta_j - l_j^-)(l_j^+ - \theta_j)}{(l_j^+ - l_j^-)^2}\right) \quad (9)$$

where,  $k$  is a scaling factor that can be used to adjust the behavior near joint limits,  $n$  represents the number of the joints in the manipulator ( $n = 7$ ),  $l_j^-$  and  $l_j^+$  are the lower and upper limit of  $j^{th}$  joint and  $\theta_j$  is the angle value of  $j^{th}$  joint. The penalization term is designed to be rapidly decreasing when the joint configuration given by a grasping pose approaches the joint limits.

By multiplying the  $M(\theta)$  with  $P(\theta)$ , a penalized manipulability can be obtained as follows:

$$C(\theta) = P(\theta)M(\theta) \quad (10)$$

In this way, the configurations that are near joint limits are penalized and make the corresponding grasping poses less likely to be selected. The comparison of selected grasping candidates with and without joint limit penalization when the object is at the same position is illustrated in Fig. 5. In Fig. 5 (a), two joints of the arm are reaching the joint limits ( $\theta_4$  and  $\theta_6$ ), which makes the arm motion unfeasible to reach the target pose. On the other hand, Fig. 5 (b) shows that by penalization, the arm joints are kept away from limits, which makes reaching motion feasible.

The change in the manipulable workspace of the shared control system before and after point cloud compensation is compared by drawing a heatmap of the manipulability (Fig. 6). The workspace is set as  $x \in [0.4m, 0.9m]$ ,  $y \in [-0.7m, 0.3m]$ , and  $z \in [0.5m, 0.8m]$  with respect to the robot base frame. Then, an object is placed in the workspace with its position changes 1cm steps at a time for a total of 150,000 positions. The GPD generates grasping candidates for each position. Using the obtained data, heat maps that contain the penalized manipulability greater than 0.05 (equation (10)) are generated. By comparing the heat map Fig. 6 (a), and Fig. 6 (b), after the point cloud compensation, the range of high manipulability grows by about 124%. This indicates the increased assistive area for the proposed shared control framework.

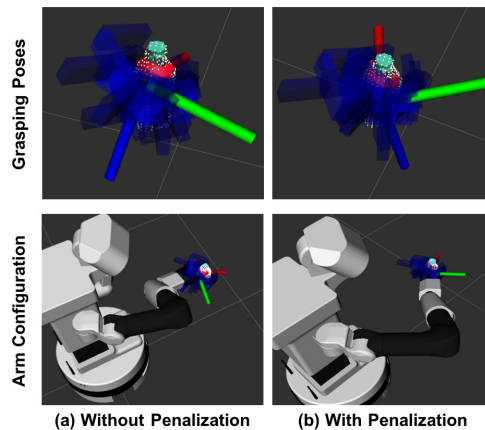


FIGURE 5. Comparison of manipulability calculation (a) without and (b) with joint limit penalization. The blue markers show the generated grasping poses, and the markers with the RGB axes are the selected grasping poses.

## C. SHARED CONTROL FRAMEWORK WITH DYNAMIC AUTHORITY ADJUSTMENT

In shared control, the distribution of authority between the operator’s input and robot execution is an important factor. This factor is generally represented as a function  $\alpha$  that represents the level of human control [37]. In this study, the  $\alpha$  is designed to be a Sigmoid function with respect to distance, as follows:

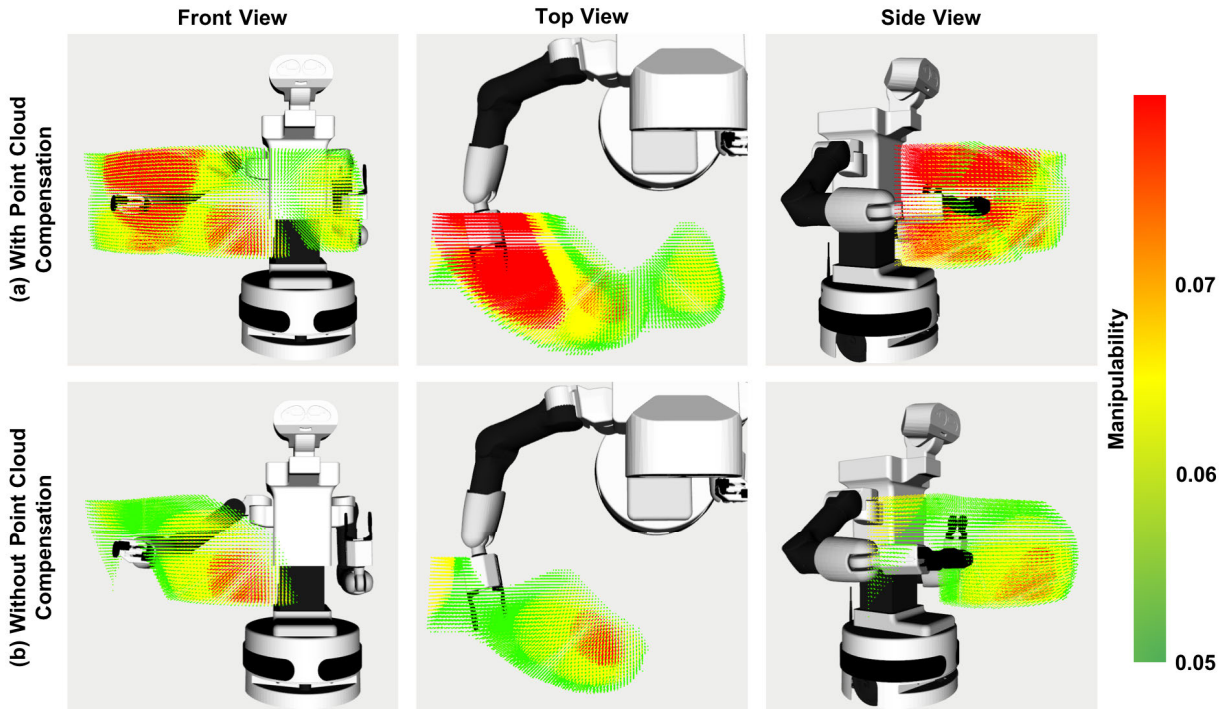
$$\alpha(t) = \frac{1}{1 - e^{-\sigma(a(t)-0.5)}} \quad (11)$$

where  $\sigma$  is a scaling factor. And  $a(t) = |P_h(t) - P_g|/d$ ,  $d = 0.3m$  is a function that maps the distance between the end-effector and the target into the range of [0.0, 1.0]. The design of the authority function  $\alpha(t)$  ensures the smooth convergence of the operator’s authority towards 0 when approaching the target object.

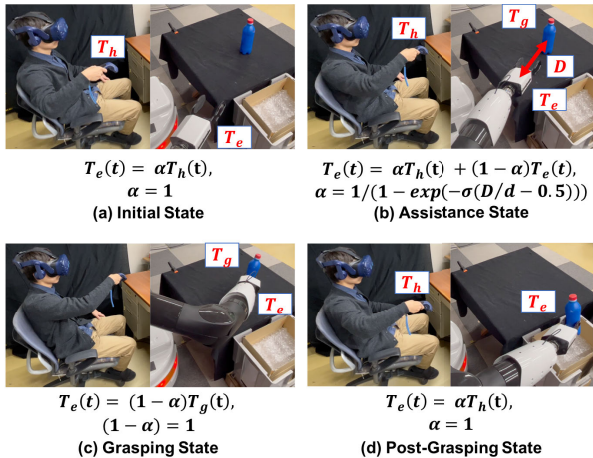
As Fig. 7 shows, in the proposed shared control framework, the whole grasping process can be divided into four states based on the value of  $\alpha$ . The  $T_e$ ,  $T_h$ ,  $T_g$ , and  $D = |P_h - P_g|$  represent the pose of the end-effector, the operator’s hand, the grasping target, and the distance to the grasping target, respectively. In Fig. 7 (a), the system is initialized and a grasping pose is generated for each object in the scene, the operator is in full control of the robot arm, and selects the target object using the VR controller. In the assistance state (Fig. 7 (b)), where  $D \in [0.12m, 0.3m]$ , the end-effector tracks the blended trajectory. In the grasping state (Fig. 7 (c)), where  $D \in [0.0m, 0.12m]$ , the robot takes full control to complete the grasp by automatic interpolation. And in the post-grasping state (Fig. 7 (d)), the operator takes back the control authority. The shared control process is illustrated in Figure 8 as motion curves.

In the assistance state (Fig. 7 (b)), the trajectory is blended following the equations listed below. For the translational trajectory:

$$p_e(t) = (1 - \alpha)p_g + \alpha p_h(t) \quad (12)$$



**FIGURE 6.** Heat map of manipulability distribution with and without point cloud compensation. After the point cloud compensation, the range of high manipulability increases by about 124%.



**FIGURE 7.** Shared control framework.

where,  $p_e(t)$ ,  $p_g$ , and  $p_h(t)$  represent the desired end-effector position at time  $t$ , the grasping target position, and the position of the human hand at time  $t$ , respectively. For rotational trajectory blending, the spherical linear interpolation (SLERP) is used:

$$R_e(t) = SLERP(R_h(t), R_g, 1 - \alpha) \quad (13)$$

where  $R_e(t)$ ,  $R_h(t)$ , and  $R_g$  are the quaternion representation of the desired end-effector orientation, human hand orientation, and orientation of the target grasping pose, respectively. The authority factor  $\alpha$  in the equation (12) and (13) is the value computed according to equation (11).

To prevent the gripper from touching the object during the adjustment process, the  $p_g$  calculated is shifted by 12cm from the target grasping pose. The system enters into the automatic grasping state (Fig. 7 (c)) either the operator presses the controller grasping trigger or the end-effector reaches the region within 12cm from the object.

#### IV. EXPERIMENTS AND DISCUSSION

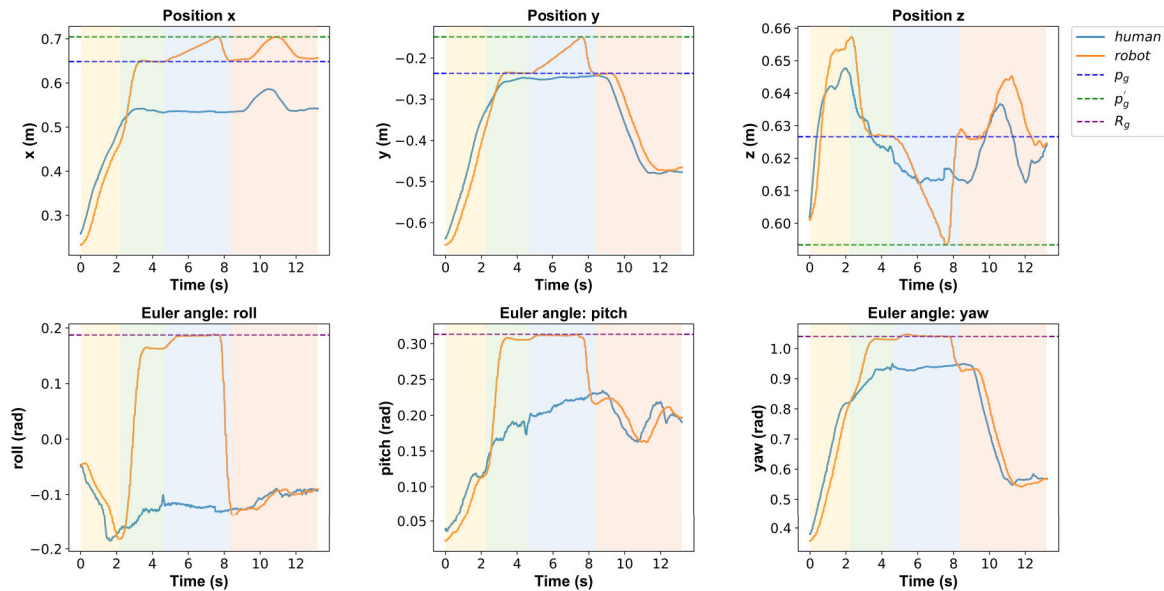
We conducted two experiments: Fig. 9 (a) a single object grasping experiment, and Fig. 9 (b) multiple objects grasping experiment to evaluate the system's performance. In the first experiment, we consider a single object to grasp, comparing the proposed shared control approach with the proposed VR device-based direct teleoperation and an LCD screen-based direct teleoperation. Then, in the second experiment, more complex multiple objects grasping is considered, comparing the shared control approach with the direct teleoperation approach. Before starting each experiment, the experimenter explained the experimental procedures for each subject. Each subject is then given time to practice the control of the teleoperation system. The experiments are carried out with 10 right-handed subjects (average age 24.4). All study participants provided informed consent, and the study was approved by the Ethics Committee of Nagoya University (No. 22-3).

##### A. EXPERIMENT #1: SINGLE OBJECT GRASPING

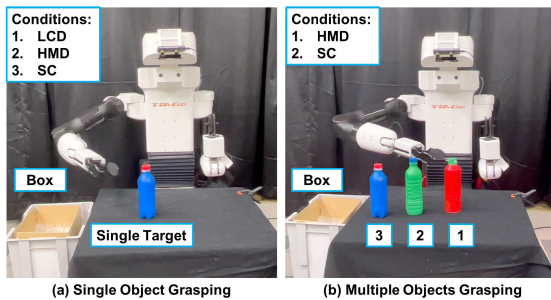
###### 1) EXPERIMENTAL SETUP AND CONDITIONS

In this experiment, only one object is considered for grasping. The experiment setup is shown in Fig. 9 (a), the target





**FIGURE 8.** Motion Curves that show the process of the shared control. The first row shows the positional changes of the human hand and the robot end-effector along the X-axis, Y-axis, and Z-axis, respectively. The second row shows the angular changes in the roll, pitch, and yaw components of both the human hand and the robot end-effector, respectively. The colors of the shaded areas indicate the different states during the shared control process (Figure 7). The light-yellow area shows the initial state where the human is in full control. The light-green area shows the assistance state where the human trajectory and robot trajectory is blended to reach the target. The light-blue area indicates the grasping state where the robot is in full control to complete the grasp. The light-pink area indicates the post-grasping state where the human is in full control to place the object inside the box.  $R_g$  indicates the orientation of the target grasping pose. Note that, to prevent the gripper from touching the object during the assistance state, the  $p_g$  calculated is shifted by 12cm from the target grasping pose  $p'_g$ .

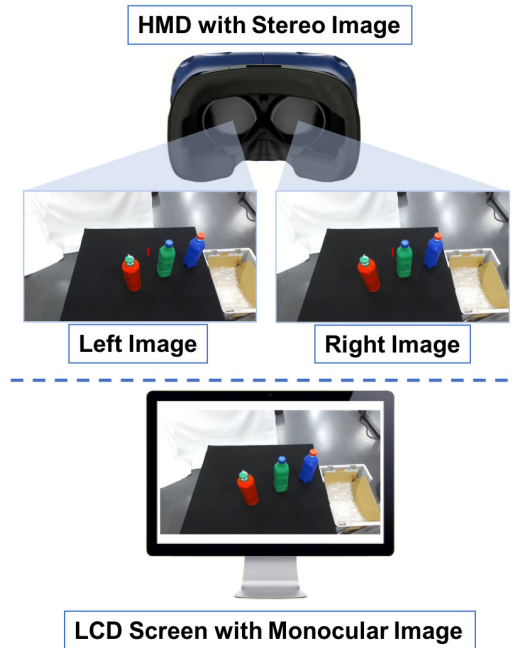


**FIGURE 9.** Grasping experiments. (a) single object grasping, and (b) multiple objects grasping.

object is a cylinder-shaped bottle ( $r=6\text{cm}$ ,  $\text{height}=20.5\text{cm}$ ) placed at  $(x=0.74\text{m}, y=-0.07\text{m})$  with respect to the robot base frame. The experimental task is to teleoperate the robot end-effector to grasp the object 5 times and place the objects in the cardboard box. During the experiment, the robot head angle is fixed to reduce the influence of other factors. The operators are required to complete the grasping under three separately given conditions:

**LCD:** A direct teleoperation, where operators control the robot end-effector by VR controller, and receive visual feedback from an LCD screen.

**HMD:** A direct teleoperation with VR display. The operators receive visual feedback from the Head Mounted Display (HMD) which presents stereo vision for enhanced spatial perception.



**FIGURE 10.** Comparison of the given visual feedback, the HMD and the LCD. The above shows the HMD with stereo-image feedback, and the number is displayed to indicate the user-selected target. The bottom shows the LCD screen with monocular image feedback.

**SC:** The proposed shared control framework, in addition to stereo vision enhanced visual feedback, the operators receive the grasping assistance for automatic



alignment to feasible grasping poses and automatic grasping completion.

The comparison of the visual feedback is shown in Figure 10. The HMD is capable of displaying stereo vision to the operator, in which the left and right images are captured through the stereo camera (Stereo Labs, ZED mini). On the other hand, the LCD screen displays the monocular image to the operator.

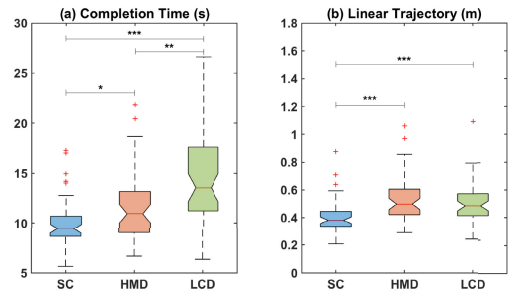
The force applied to the object is controlled through the inner position controller provided by the parallel gripper mounted on the Tiago wrist. When the operator pushes the trigger of the VR controller to close the gripper, the position controller controls the position of the two gripper fingers to close completely without the gap. The maximum force applied to the object is set to 50N to guarantee the stability of the grasp. The grasping poses generated by the GPD ensure force closure grasps will be formed with respect to the target objects. Furthermore, the target objects considered in the current experiments are rigid cylinder-shaped objects, with which we have observed stable grasps.

In the experiment, the grasping procedure is as follows: (1) the operators are asked to first place their hand in the home position (Fig. 7 (a)), (2) then move the hand for grasping and placing (Fig. 7 (b), (c), (d)), and (3) finally move back to the home position (Fig.7 (a)). The task completion time and linear trajectories are recorded to compare the operators' performance under each condition. The experiment starts when the first time the robot end-effector appears on the display, and task completion time records the time duration of each grasping starting from the robot end-effector appearing on the HMD until the object is placed in the box.

## 2) RESULTS AND DISCUSSION

To compare different metrics, one-way ANOVA tests are performed on the data. Fig. 11 (a) shows the box plot of the task completion time. Note that, “\*\*\*\*”, “\*\*\*”, and “\*” in the figure represent for  $p < 0.001$ ,  $p < 0.01$ , and  $p < 0.05$ , respectively. The median value of task completion time is 9.39s, 10.94s, and 13.56s for the condition of SC, HMD, and LCD, respectively. The one-way ANOVA test shows a statistical significance in the task completion time across the conditions between SC vs. HMD ( $p < 0.05$ ), SC vs. LCD ( $p < 0.001$ ), and HMD vs. LCD ( $p < 0.01$ ). The results (SC vs. HMD, SC vs. LCD) indicate that for a single object grasping, the proposed shared control framework could enhance the operator's performance regarding reduced task completion time. This also indicates that the grasping task is made easier by integrating robot automation. The result of HMD vs. LCD indicates that, by using stereo vision for visual feedback, operators have the better spatial perception that can accelerate the process of grasping pose alignment.

Fig. 11 (b) shows the box plot of the linear trajectory of the human hand. The one-way ANOVA test reveals a statistical significance in the task linear trajectory across the conditions between SC vs. HMD ( $p < 0.001$ ), and SC vs. LCD ( $p < 0.001$ ). The median value of linear trajectory is 0.38m,



**FIGURE 11. Experimental result of single object grasping. (a) Task Completion time, and (b) Linear trajectory.**

0.49m, and 0.48m for the condition of SC, HMD, and LCD, respectively. This result indicates that with shared control human hands travel shorter distances compared with direct teleoperation. The shorter moving distance could reduce the workload of the operators during the grasping process. On the other hand, statistical significance is not observed in the condition of HMD vs. LCD. Since both condition HMD and LCD are based on direct teleoperation, where the follower end-effector constantly tracks operators' hands, the operators must reach the full distance towards the targeted object. This additional movement potentially increases the workload during teleoperation.

## B. EXPERIMENT #2: MULTIPLE OBJECTS GRASPING

### 1) EXPERIMENTAL SETUP AND CONDITIONS

In this experiment, multiple objects are considered for grasping. The experiment setup is shown in Fig. 9 (b), the target objects are cylinder-shaped bottles with different colors: #1 red ( $r=6.5\text{cm}$ ,  $h=21.5\text{cm}$ ), #2 green ( $r=6.5\text{cm}$ ,  $h=22.5\text{cm}$ ), #3 blue ( $r=6\text{cm}$ ,  $h=20.5\text{cm}$ ), and placed at different places ( $x=0.70\text{m}$ ,  $y=0.07\text{m}$ ), ( $x=0.74\text{m}$ ,  $y=-0.07\text{m}$ ), ( $x=0.76\text{m}$ ,  $y=-0.23\text{m}$ ) in the robot base frame, respectively. The experimental task is to teleoperate the robot end-effector to grasp the 3 objects and place the objects to the cardboard box. Each operator is asked to perform the task 3 times. During the experiment, the robot head is fixed to a predefined angle to reduce the influence of other factors. The operators are required to complete the grasping task under two separately given conditions: (1) **SC**, and (2) **HMD**. The details of each condition are described in experiment #1.

The process to generate grasping poses for different target objects follows Algorithm 1. The algorithm first takes the environment point cloud which contains target objects as an input. It then segments this point cloud to acquire individual clusters for each object, each representing a partially observed point cloud of the object. By applying template matching for each cluster, the compensated object point cloud which contains positional information can be obtained. The grasping poses for each object can be acquired by running the Grasping Pose Detection algorithm to the compensated point cloud. These poses are filtered by manipulability score to yield feasible grasping poses.

In the experiment, operators teleoperate the robot to grasp each object following the procedures in experiment #1. The

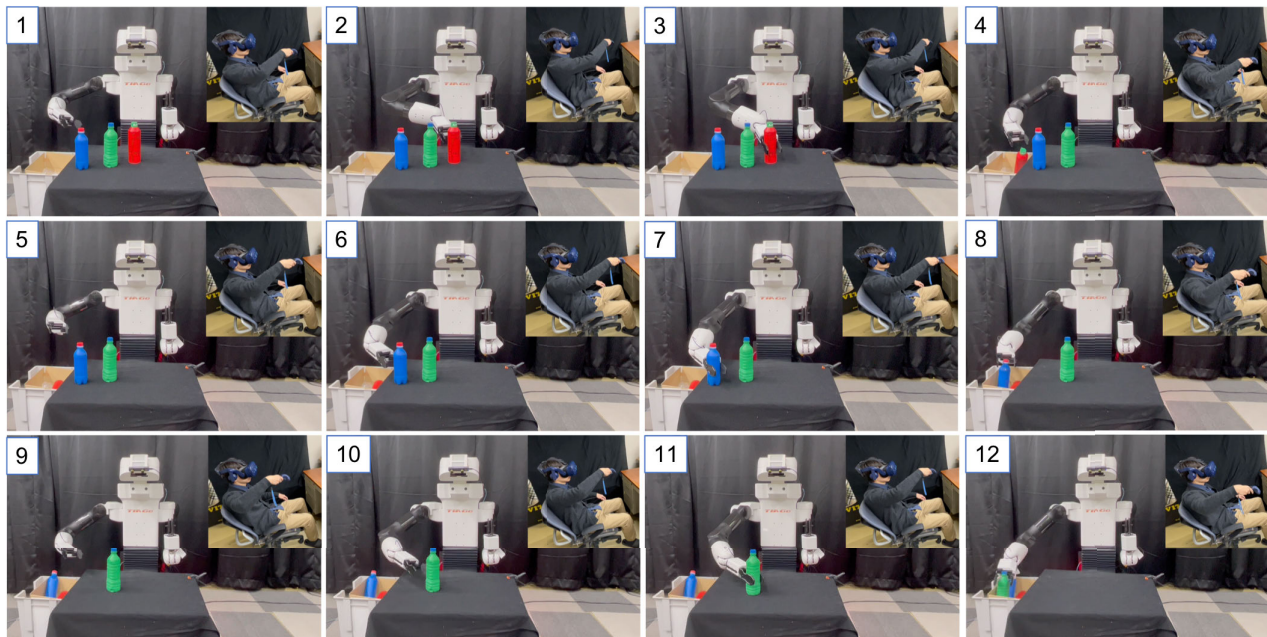


FIGURE 12. Screenshots that show the process of multiple objects grasping with shared control.

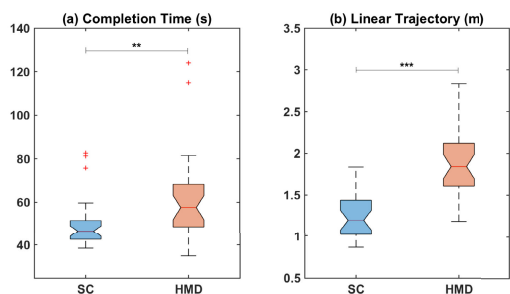


FIGURE 13. Experimental result of multiple objects grasping. (a) Task Completion time, and (b) Linear trajectory.

shared control for grasping each target object is achieved following the four states described in Section III, subsection C and depicted in Figure 7. Operators are asked to grasp the target object following the order of object #1, object #3, and object #2. After grasping each object, operators return to the home position and select the next target object by using the trackpad of the VR controller. The ID of the selected object is displayed on the HMD. The task completion time is measured starting from the first time the robot end-effector appears on the HMD until the last object is placed in the box. The linear trajectories track the whole grasping process. The screenshots given by Figure 12 illustrate the process of multiple objects grasping with shared control.

2) RESULTS AND DISCUSSION

Fig. 13 (a) shows the box plot of the task completion time. The one-way ANOVA test shows a statistical significance in the task completion time between the condition of SC and HMD ( $p < 0.01$ ). The median value of task completion time is 46.52s, and 57.54s, for the condition

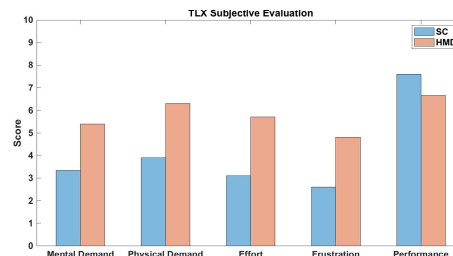


FIGURE 14. NASA TLX evaluation for multiple objects grasping.

of SC, and HMD, respectively. The statistical significance between SC and HMD reveals that even for more complex multiple objects grasping, in addition to the enhanced visual perception, the proposed shared control can enhance the grasping performance by means of auto-alignment of the grasping pose and auto-completion.

Fig. 13 (b) shows the box plot of the linear trajectory of the human hand. A statistical significance is revealed by the ANOVA test for the condition of SC and HMD ( $p < 0.001$ ). The median values are 1.196m, and 1.841m for SC and HMD, respectively. The result is similar to single object grasping, by using the proposed shared control, human hands travel shorter distances and hence decrease the workload. Note that, none of the subjects have previous experience with the operation of the teleoperation system. In addition, the Interquartile Range (IQR) of the completion time is 8.92s, and 19.64s for SC and HMD, respectively. The IQR of the linear trajectories is 0.420m, and 0.513m for SC and HMD, respectively. The IQR of the box plots suggests that the data of the shared control condition has less dispersion. This could imply that despite the operator’s experience, the shared control approach can assist the operators to execute the task

in a more stable way. Moreover, the NASA TLX subjective evaluation (Fig. 14) shows less mental demand, physical demand, effort, and frustration during the experiment #2. This result further verifies the previous results which indicate the reduced workload when using the proposed shared control. As a result, the subjective measurement of the shared control's performance is greater than the direct teleoperation.

## V. CONCLUSION

This paper proposes a shared control framework for grasping assistance in teleoperation. The presented shared control framework is built upon an intuitive VR device-based direct teleoperation system. In which, human hand and head motions are captured by HTC VR devices, and then mapped to an anthropomorphic robotic manipulator (Pal Robotics, Tiago) through an inverse kinematics solver with null-space resolution.

In the shared control framework, a template matching-based point cloud compensation is performed for multi-angle grasping pose generation. Then a joint limit penalized manipulability analysis is applied to the generated grasping poses to acquire the most feasible pose candidates. By applying for the object point cloud compensation, the assistive area with high manipulability is increased about 124%. The shared control is then achieved by dynamic authority adjustment-based trajectory blending. Two grasping experiments are carried out for system evaluation. For both experiments, enhanced performances are observed by means of faster task completion time, and reduced linear trajectories. The NASA TLX subjective evaluation also showed a reduced workload for multiple object-grasping tasks.

The current point cloud compensation is targeted for symmetrically shaped objects with controlled lighting conditions. In the future to further improve the system capability, the point cloud compensation for the asymmetrically shaped object given more complex lighting conditions and backgrounds is going to be carried out. The future approach will be guided by methodologies reported by computer vision researchers [38], [39], [40]. Additionally, the use of tactile sensing technology can be explored in the future to control the force applied to the objects and enable the stable grasping of deformable objects or complex shaped objects such as fruits. Human comfortableness is also an important factor in designing an effective teleoperation system, and this aspect will be explored in the future. The collision avoidance method can be integrated into the current system for task execution in a more complex and cluttered environment. In addition, the assistance in the post-grasping phase such as placing an object in the desired location, or handover an object to a human can be explored.

## REFERENCES

- [1] R. C. Goertz, "Manipulator systems developed at ANL," in *Proc. 12th Conf. Remote Syst. Technol.*, 1964, pp. 117–136.
- [2] Y. Zhu, T. Aoyama, and Y. Hasegawa, "Enhancing the transparency by onomatopoeia for passivity-based time-delayed teleoperation," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 2981–2986, Apr. 2020.
- [3] D. A. Lawrence, "Stability and transparency in bilateral teleoperation," *IEEE Trans. Robot. Autom.*, vol. 9, no. 5, pp. 624–637, Oct. 1993.
- [4] M. Selvaggio, M. Cognetti, S. Nikolaidis, S. Ivaldi, and B. Siciliano, "Autonomy in physical human-robot interaction: A brief survey," *IEEE Robot. Autom. Lett.*, vol. 6, no. 4, pp. 7989–7996, Oct. 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9501975/>
- [5] D. Rakita, B. Mutlu, M. Gleicher, and L. M. Hiatt, "Shared control-based bimanual robot manipulation," *Sci. Robot.*, vol. 4, no. 30, May 2019, Art. no. eaaw0955.
- [6] R. Balachandran, H. Mishra, M. Cappelli, B. Weber, C. Secchi, C. Ott, and A. Albu-Schaeffer, "Adaptive authority allocation in shared control of robots using Bayesian filters," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 11298–11304.
- [7] M. Selvaggio, A. M. E. Ghalamzan, R. Moccia, F. Ficuciello, and B. Siciliano, "Haptic-guided shared control for needle grasping optimization in minimally invasive robotic surgery," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 3617–3623.
- [8] F. Abi-Farraj, N. Pedemonte, and P. R. Giordano, "A visual-based shared control architecture for remote telemanipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 4266–4273.
- [9] F. Huang, X. Chen, Z. Chen, and Y. Pan, "A novel SMMS teleoperation control framework for multiple mobile agents with obstacles avoidance by leader selection," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 3, pp. 1517–1529, Mar. 2023.
- [10] G. Quere, A. Hagenhuber, M. Iskandar, S. Bustamante, D. Leidner, F. Stulp, and J. Vogel, "Shared control templates for assistive robotics," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 1956–1962.
- [11] M. Selvaggio, F. Abi-Farraj, C. Pacchierotti, P. R. Giordano, and B. Siciliano, "Haptic-based shared-control methods for a dual-arm system," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 4249–4256, Oct. 2018.
- [12] P. Kazanzides, B. P. Vagvolgyi, W. Pryor, A. Deguet, S. Leonard, and L. L. Whitcomb, "Teleoperation and visualization interfaces for remote intervention in space," *Front. Robot. AI*, vol. 8, Dec. 2021, Art. no. 747917.
- [13] F. Abi-Farraj, C. Pacchierotti, and P. R. Giordano, "User evaluation of a haptic-enabled shared-control approach for robotic telemanipulation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 1–9.
- [14] T. Lin, A. U. Krishnan, and Z. Li, "Shared autonomous interface for reducing physical effort in robot teleoperation via human motion mapping," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 9157–9163.
- [15] F. Abi-Farraj, C. Pacchierotti, O. Arenz, G. Neumann, and P. R. Giordano, "A haptic shared-control architecture for guided multi-target robotic grasping," *IEEE Trans. Haptics*, vol. 13, no. 2, pp. 270–285, Apr. 2020.
- [16] A. ten Pas, M. Gualtieri, K. Saenko, and R. Platt, "Grasp pose detection in point clouds," *Int. J. Robot. Res.*, vol. 36, nos. 13–14, pp. 1455–1473, Dec. 2017.
- [17] E. A. M. Ghalamzan, F. Abi-Farraj, P. R. Giordano, and R. Stolkin, "Human-in-the-loop optimisation: Mixed initiative grasping for optimally facilitating post-grasp manipulative actions," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 3386–3393.
- [18] Y. Xu, C. Ding, X. Shu, K. Gui, Y. Bezudnova, X. Sheng, and D. Zhang, "Shared control of a robotic arm using non-invasive brain-computer interface and computer vision guidance," *Robot. Auto. Syst.*, vol. 115, pp. 121–129, May 2019.
- [19] M. Laghi, L. Raiano, F. Amadio, F. Rollo, A. Zunino, and A. Ajoudani, "A target-guided telemanipulation architecture for assisted grasping," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 8759–8766, Oct. 2022.
- [20] M. Bowman and X. Zhang, "Dimension-specific shared autonomy for handling disagreement in telemanipulation," *IEEE Robot. Autom. Lett.*, vol. 8, no. 3, pp. 1415–1422, Mar. 2023.
- [21] S. Jauhri, J. Peters, and G. Chalvatzaki, "Robot learning of mobile manipulation with reachability behavior priors," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 8399–8406, Jul. 2022.
- [22] T. Yamamoto, K. Terada, A. Ochiai, F. Saito, Y. Asahara, and K. Murase, "Development of human support robot as the research platform of a domestic mobile manipulator," *ROBOMECH J.*, vol. 6, no. 1, pp. 1–15, Dec. 2019.
- [23] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 5628–5635.



[24] D. Lee, D. Ko, W. K. Chung, and K. Kim, "Maximal manipulation framework using quadratic programming for a teleoperated robotic system with articulated bodies," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2022, pp. 9339–9345.

[25] P. Stotko, S. Krumpfen, M. Schwarz, C. Lenz, S. Behnke, R. Klein, and M. Weinmann, "A VR system for immersive teleoperation and live exploration with a mobile robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2019, pp. 3630–3637.

[26] D. G. Macharet and D. A. Florencio, "A collaborative control system for telepresence robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2012, pp. 5105–5111.

[27] D. Sun, A. Kiselev, Q. Liao, T. Stoyanov, and A. Loutfi, "A new mixed-reality-based teleoperation system for telepresence and maneuverability enhancement," *IEEE Trans. Hum.-Mach. Syst.*, vol. 50, no. 1, pp. 55–67, Feb. 2020.

[28] F. Flacco, A. De Luca, and O. Khatib, "Control of redundant robots under hard joint constraints: Saturation in the null space," *IEEE Trans. Robot.*, vol. 31, no. 3, pp. 637–654, Jun. 2015.

[29] P. Liu, D. Tateo, H. Bou-Ammar, and J. Peters, "Efficient and reactive planning for high speed robot air hockey," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2021, pp. 586–593.

[30] J. Nakanishi, R. Cory, M. Mistry, J. Peters, and S. Schaal, "Operational space control: A theoretical and empirical comparison," *Int. J. Robot. Res.*, vol. 27, no. 6, pp. 737–757, Jun. 2008.

[31] J. Nakanishi, S. Itadera, T. Aoyama, and Y. Hasegawa, "Towards the development of an intuitive teleoperation system for human support robot using a VR device," *Adv. Robot.*, vol. 34, no. 19, pp. 1239–1253, Oct. 2020.

[32] J. Carpentier, G. Saurel, G. Buondonno, J. Mirabel, F. Lamiroux, O. Stasse, and N. Mansard, "The Pinocchio C++ library—A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Jan. 2019, pp. 614–619.

[33] R. B. Rusu and S. Cousins, "3D is here: Point cloud library (PCL)," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 1–4.

[34] K. G. Derpanis, "Overview of the RANSAC algorithm," *Image Rochester NY*, vol. 4, no. 1, pp. 2–3, 2010.

[35] T. Yoshikawa, "Manipulability of robotic mechanisms," *Int. J. Robot. Res.*, vol. 4, no. 2, pp. 3–9, Jun. 1985.

[36] N. Vahrenkamp, T. Asfour, G. Metta, G. Sandini, and R. Dillmann, "Manipulability analysis," in *Proc. 12th IEEE-RAS Int. Conf. Humanoid Robots (Humanoids)*, Nov. 2012, pp. 568–573.

[37] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *Int. J. Robot. Res.*, vol. 32, no. 7, pp. 790–805, Jun. 2013.

[38] D. Lee, O. C. Hamsici, S. Feng, P. Sharma, and T. Gernoth, "DeepPRO: Deep partial point cloud registration of objects," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 5663–5672.

[39] Y. Li and T. Harada, "Leopard: Learning partial point cloud matching in rigid and deformable scenes," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5544–5554.

[40] R. Vock, A. Dieckmann, S. Ochmann, and R. Klein, "Fast template matching and pose estimation in 3D point clouds," *Comput. Graph.*, vol. 79, pp. 36–45, Apr. 2019.



**YAONAN ZHU** (Member, IEEE) received the B.S. degree from Northeastern University, China, in 2015, the M.S. degree from the University of Science and Technology of China, in 2018, and the Ph.D. degree in robotics from Nagoya University, Japan, in 2021. He is currently a designated Assistant Professor at Nagoya University. His research interests include telerobotics and teleoperation, haptics and haptic interfaces, human-centered robotics, and wearable robotics.

He received the Best Paper Award Finalist at the 2022 IEEE International Conference on Cyborg and Bionic Systems (CBS).



**BINGHENG JIANG** received the B.S. degree from Shanghai Jiao Tong University, China, in 2019, and the M.S. degree in robotics from Nagoya University, Japan, in 2022. He is currently a Software Engineer at IAS BU, HUAWEI, China. His research interests include robot teleoperation, shared control, and VR interface.



**QIBIN CHEN** received the B.S. degree from Northeastern University, China, in 2021. He is currently pursuing the master's degree with Nagoya University, Japan. His research interests include telerobotics and teleoperation, assistive robotics, and surgical robotics.



**TADAYOSHI AOYAMA** (Member, IEEE) received the B.E. degree in mechanical engineering, the M.E. degree in mechanical science engineering, and the Ph.D. degree in micro-nano systems engineering from Nagoya University, Nagoya, Japan, in 2007, 2009, and 2012, respectively. He was an Assistant Professor at Hiroshima University, Japan, from 2012 to 2017, and at Nagoya University, from 2017 to 2019. He was also a PRESTO Researcher at JST, from 2018 to 2022.

He is currently an Associate Professor at the Department of Micro-Nano Mechanical Science and Engineering, Nagoya University. His research interests include macro-micro interaction, VR/AR and human interface, AI-based assistive technology, micro-manipulation, and medical robotics. He received the Best Paper Award from the 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM2019) and the IEEE International Symposium on Micro-NanoMechatronics and Human Science (MHS2017, MHS2019, and MHS2022).



**YASUHISA HASEGAWA** (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in robotics from Nagoya University, Nagoya, Japan, in 1994, 1996, and 2001, respectively. From 1996 to 1998, he was with Mitsubishi Heavy Industries Ltd., Japan. He joined Nagoya University, in 1998. He moved to Gifu University, in 2003. From 2004 to 2014, he was with the University of Tsukuba. Since 2014, he has been with Nagoya University, where he is currently

a Professor with the Department of Micro-Nano Systems Engineering. His research interests include motion-assistive systems, teleoperation for manipulation, and surgical support robot.

...