A Stable Guidance Method for Teleoperation-based Robot Learning from Demonstration

Weiyong Si†, Zhehao Jin†, Zhenyu Lu, Ning Wang and Chenguang Yang*

***Abstract*—Robot learning from teleoperation-based demonstration has been proven to be an effective method for robot skills learning. However, the human demonstration process is mentally and physically challenging due to the dexterous and continuous interaction control. Especially, for contact-rich tasks human operators need to control the pose and contact force between the end-effector and the environment, only based on visual and haptic feedback. To tackle the problem, this work studies a stable guidance method for teleoperation-based demonstration to improve the experience and performance of human demonstration and reduce the workload of human operators. The guidance design is based on a dynamic system-based imitation learning model, which is achieved by human demonstration. The virtual guidance force will decrease as the virtual guidance's iterative learning. The efficiency of human demonstration is improved, and human operators' physical and mental workload is reduced. Two typical experiments, robot-assisted polishing and robot-assisted ultrasound scanning, are conducted to validate the performance of the guidance method.**

I. Introduction

Telerobotics has been studied since 1950 and various multimodal feedback-based teleoperation systems have been developed [1]. Teleoperation provides a solution for remote operations for dangerous environments and tasks, so telerobotics was used in various fields, such as the medical field, space exploration, nuclear industry and rehabilitation [2], [3]. Nowadays, robot learning through teleoperation-based demonstration has attained much attention because teleoperation provides a human-in-the-loop mechanism for online demonstration, interaction and correction [1], [4]. One of the main components in the teleoperation-based demonstration is the teleoperation interface, also named the human-robot interface.

The safety and intuitiveness of the human-robot interface are challenging. To address these challenges, various methods have been proposed, such as the shared control method [5], [6], virtual guidance [7], and immersive teleoperation devices [8]. The shared control uses human-robot collaboration to complete challenging tasks by reducing the workload of human operators [9]. In the shared control framework, humans only focus on intelligent and decision-making tasks, such as task decisions, while robots are responsible for certain repetitive tasks. In addition, various human-robot interactive interfaces based on state-of-the-art hardware devices have been developed to benefit teleoperation, such as virtual reality (VR) [10], augmented reality (AR), haptic devices, etc., which can provide immersive telepresence [8]. Furthermore, one of the methods for teleoperation is virtual guidance for dexterous manipulation [11].

The virtual guidance can constrain robots' motion to satisfy the task-relevant requirements, which could improve safety and control performance [12]. This has been employed in medical applications and flexible manufacturing. In [13], a library of virtual guiding fixtures has been studied to assist multiple tasks by providing intuitive haptic guidance, which proves this method could improve performance in terms of efficiency, human workload and control accuracy. However, the existing guidance methods usually require dedicated designs for a specific task, hence they are less generalised to new tasks.

The virtual guidance methods have been proposed in the existing papers [7], [13]–[15]. Virtual fixtures for bilateral teleoperation have been proposed for medical applications and manufacturing. For example, the interactive generation of virtual guidance method for manufacturing tasks in unstructured environments is studied in [14]. In addition, vision-based virtual guidance (VG) generation methods are studied for bilateral teleoperation. Furthermore, VG can also be employed in multiple-user teleoperations for robotic surgical training [16]. Learning from demonstration-based methods for VG design has been proposed [17]. The imitation learning method was used for human intention prediction and shared control [18]. Various learning from demonstration (LfD) methods and frameworks have been developed for human-robot skill transfer, specifically for the contact-rich manipulation tasks [4], [19]. The LfD method offers advantages in learning efficiency and generalization, making it suitable for learning motion, force, and stiffness skills [20]. Additionally, the LfD method has been investigated for applications in physical human-robot interaction, among other areas.

The VG has been studied to assist the human demonstration of human-robot skills transfer [21]. However, this VG is only considered in the kinesthetic guidance. In [22], the author studied the vision-based VG for dexterous manipulation. Although the above work has been proven ineffective in teleoperation and human-robot interaction, designing the VG needs more time and effort. In addition, most of the existing VG methods assume the design rules are designed in advance, and the design by humans is optimal.

† W. Si and Z. Jin contribute equally.

W. Si is with the School of Computer Science and Electronic Engineering at the University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK. Z. Jin is with the College of Information Engineering, Zhejiang University of Technology, Hangzhou, 310032, China. Z. Lu and N. Wang are with Bristol Robotics Lab at the University of the West of England, Bristol, BS16 1QY, UK. C. Yang is the Department of Computer Science, University of Liverpool, Liverpool, L69 3BX, UK.

* Corresponding author is C. Yang (Email: cyang@ieee.org)

This paper aims to benefit the design of a virtual guidance method for teleoperation and robot learning from teleoperation-based demonstration. We propose a novel VG design method for tele-manipulation, especially for safety-critical medical examination tasks, such as ultrasound scanning. We studied the human demonstration method to attain rough skills, based on which we got a neural energy function with a unique minimum (NEUM) [23]. The NEUM can plan a trajectory from the high energy points to the low energy points, which can finally converge to the global minimum point [23]. Comparing the existing energy VG method, such as the artificial potential field-based virtual fixture [24], the NEUM-based guidance can guarantee convergence globally. We evaluate the effectiveness of the proposed method through robot-assisted ultrasound scanning tasks. The main contributions of this paper are summarized below:

- The VG design investigates a NEUM-based imitation learning model which is built based on human demonstration, for human-robot skill transfer. Comparing the existing VG method, the proposed method could only need demonstration to build the guidance, without complex design and parameters modification.
- In addition, NEUM-based VG can also guarantee the convergence of the desired VG. Learned skills provide guidance for human teachers when doing teleoperation-based demonstrations. This VG method has been evaluated through robot-assisted ultrasound scanning and polishing tasks.

In the following section II, the bilateral teleoperation control system is designed, and the proposed VG method is also presented. The experiment system setup is given in Section III. In Section IV, we conduct experiments to evaluate the effectiveness and performance of the proposed method. The results are given in Section V. In Section VI, we discuss the results and conclude this paper and potential work in the future.

II. METHODOLOGY

A. Bilateral teleoperation control design

For bilateral teleoperation, the dynamics of the robot system in Cartesian space are given [25],

$$A_L(q_L)\ddot{x}_L + B_L(q_L,\dot{q}_L)\dot{x}_L + G_L(q_L) = u_L + f_h \tag{1}$$

$$A_F(q_F)\ddot{x}_F + B_F(q_F,\dot{q}_F)\dot{x}_F + G_F(q_F) = u_F + f_e \tag{2}$$

where $A_L(q_L)$ and $A_F(q_F)$ are the inertia matrix, $B_L(q_L,\dot{q}_L)$ and $B_F(q_F,\dot{q}_F)$ are the Coriolis and centrifugal terms, and $G_L(q_L)$ and $G_F(q_F)$ represent the gravity of the leader and follower robots[1] respectively. u_L and f_h are the control input and operator force of the leader robot. q_L and $\dot{q}_L$ are the joint position and velocity of leader robot. The $\dot{x}_L$ and $\ddot{x}_L$ are the velocity and acceleration of the leader robot in Cartesian space. u_F and f_e are the control and interaction force of the follower robot. q_F and $\dot{q}_F$ are the joint position and velocity of the follower robot, the $\dot{x}_F$ and $\ddot{x}_F$ represent the velocity and acceleration of the follower robot.

Impedance control builds the models of the relationship between the robot and the environment as a mass-spring-damper system [25]. We designed the impedance controller in task space for the follower robot,

$$u_F = K_F(x_L - x_F) - D_F\dot{x}_F \tag{3}$$

where u_F is the control command of follower robot, K_F is the stiffness matrix, the D_F represents the damping matrix. x_L and x_F are the positions of the leader and follower robots, respectively.

The force feedback is designed to reflect the interaction force between the follower robot and its environment. The controller of the leader robot is designed as follows,

$$u_L = -K_L f_e - D_L\dot{x}_L \tag{4}$$

where u_L is the control input of the leader robot, D_L is the damping matrix. f_e is the interaction force between the follower robot and environment, K_L is the scaling parameter. One advantage of bilateral teleoperation is that force feedback for the human operator benefits the human-robot skill transfer. To reflect the contact force on the haptic device side, we design the following controller,

$$f_m = -K_L f_e \tag{5}$$

where K_L is the scaling parameter to transform the contact force on the end-effector side to the haptic side.

B. NEUM and dynamic system learning based on the human demonstration

To build the NEUM, we attain the demonstration data set by teleoperation. $D = \{x, \dot{x}\}_{t=0:t_n}^{T=0:N}$ is the demonstration data set, where x and $\dot{x}$ are the position and velocity of end-effector (tools, including the polishing tool and ultrasound probe in the Cartesian space), respectively.

We developed the NEUM for human skills encoding [23], which has various advantages. The NEUM has the form,

$$\begin{cases} V(x) = V_1(x) - V_1(0) + V_2(x) \\ V_1(x) = \omega^T f(x) \\ f(x) = \left[f_1(x), \cdots, f_k(x), \cdots, f_{d_H}(x)\right]^T \\ f_k(x) = \varrho(a_k^T z(x) + b_k) \\ z(x) = \left[\|x\|_2^{1+\epsilon} \quad \|x\|_2^{\epsilon} x^T\right]^T \\ V_2(x) = \alpha x^T x \end{cases} \tag{6}$$

where $V_1(x)$ is a neural network with weight parameter $\omega \in R^{d_H}$ and feature $f(x) : R^{d_x} \rightarrow R^{d_H}$. The activation function is designed as $\varrho(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}}$. Function $z(x) : R^{d_x} \rightarrow R^{d_x+1}$ is a manually designed encoder, $a_k \in R^{d_x+1}$ and $b_k \in R$ are feature parameters, ϵ, $\alpha \in R_{++}$ are positive scalars. Function $V_2(x)$ is used to ensure the radially unbounded property of the $V(x)$. It has been shown in our previous work [23] that if the following constraints hold,

$$\begin{cases} a_{k,1} > 0 \\ a_{k,1}^2 - \sum_{i=2}^{d_x+1} a_{k,i}^2 > 0, \forall k \in [1, \cdots, d_H] \\ \omega_k > 0 \end{cases}, \tag{7}$$

[1] Note the leader robot refers to the teleoperation device, Touch X, and the follower robot refers to the Franka robot manipulator.

the NEUM $V(x)$ has a unique minimum located at $x=0$, and simultaneously be positive-definite, radially unbounded and continuously differentiable.

The NEUM can then be learned by solving the following optimization problem

$$\min_{\Theta} J(\Theta) \tag{8}$$

$$\text{subject to (7)}$$

where $\Theta = \{a_k, b_k, \omega\}_{k=[1,\cdots,d_H]}$ is the collection of the learned parameters, and the objective function $J(\Theta)$ is defined as follows

$$\begin{cases} J(\Theta) = \sum\limits_{t,n} j_{t,n}(\Theta) + L_2\|\Theta\|_2^2 \\ j_{t,n}(\Theta) = tanh(\beta \bar{j}_{t,n}) = \frac{e^{\beta \bar{j}_{t,n}} - e^{-\beta \bar{j}_{t,n}}}{e^{\beta \bar{j}_{t,n}} + e^{-\beta \bar{j}_{t,n}}} \\ \bar{j}_{t,n}(\Theta) = \frac{\dot{x}^T \frac{\partial V(x_{t,n},\Theta)}{\partial x_{t,n}}}{\|\dot{x}\|_2 \| \frac{\partial V(x_{t,n},\Theta)}{\partial x_{t,n}}\|_2} \end{cases}, \tag{9}$$

Using the learned NEUM, we can then learn a globally stable autonomous dynamic system (ADS) to encode the human demonstration preferences. The stable ADS is formulated as $\dot{x} = g(x) = o(x) + u$, where $o(x)$ is the original ADS function, which can be learned via any regression algorithm on the demonstration data set D. The correct term u can be obtained by online solving the following QP problem,

$$\min_{u} u^T u \tag{10}$$

subject to

$$(o(x)+u)^T \frac{\partial V(x)}{\partial x} \leqslant -\rho(x) \tag{11}$$

The optimization problem admits a closed-form solution, which is provided as follows:

$$u = -\frac{ReLu\left(o(x)^T \frac{\partial V(x)}{\partial x} + \rho(x)\right)}{\left\|\frac{\partial V(x)}{\partial x}\right\|_2^2} \frac{\partial V(x)}{\partial x} \tag{12}$$

where the activation function $ReLu(s)$ has the form

$$ReLu(s) = \begin{cases} s, s>0 \\ 0, else \end{cases} \tag{13}$$

As a result, the stable ADS is given as

$$\dot{x} = o(x) - \frac{ReLu\left(o(x)^T \frac{\partial V(x)}{\partial x} + \rho(x)\right)}{\left\|\frac{\partial V(x)}{\partial x}\right\|_2^2} \frac{\partial V(x)}{\partial x} \tag{14}$$

where $V(x)$ is given in (6).

C. Virtual guidance design

The VG is designed based on the guidance of NEUM. In this work, we only consider the force feedback, and ignore the torque feedback in orientation. Thus the virtual force is designed as the following,

$$V_{FR} = K(P_d - P_c) + D(V_d - V_d) \tag{15}$$

where $V_{FR} \in R^{3\times1}$ is the virtual force, $P_d \in R^{3\times1}$ and $V_d \in R^{3\times1}$ are the desired position and velocity respectively.

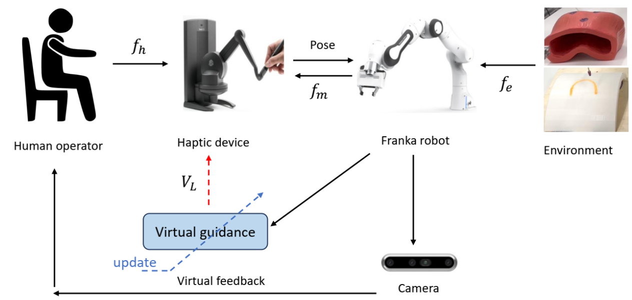

Fig. 1. The block diagram of bilateral teleoperation enhanced by stable VG.

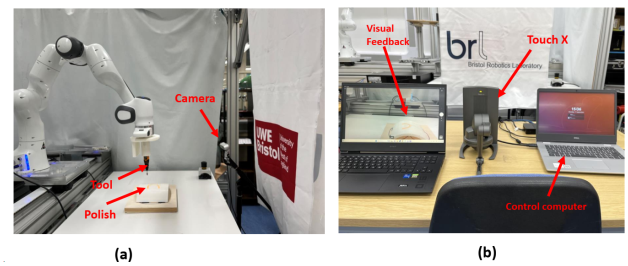

Fig. 2. The setup of the bilateral teleoperation platform.

$P_c \in R^{3\times1}$ and $V_c \in R^{3\times1}$ are the desired position and velocity respectively where $K \in R^{3\times3}$ and $D \in R^{3\times3}$ are the spring and damping terms for the virtual system respectively.

To transmit the virtual force from the follower robot side to the leader device side,

$$V_{LB} = {}^{LB}_{FB} R V_{FB} \tag{16}$$

where ${}^{LB}_{FB}R$ is the rotation matrix from the base frame of the leader device to the base frame of the follower robot. V_{FR} and the V_{LR} are the virtual force in follower robot side and leader robot side respectively.

$$F_L = f_c + V_{LB} \tag{17}$$

$$f_c = K_{sc} f_{FC} \tag{18}$$

where $f_{FC} \in R^{3\times1}$ is the contact force between the end-effector and environment on follower robot side, $K_{sc} \in R^{3\times1}$ is the scaling parameter for force feedback, $f_c \in R^{3\times1}$ is the force feedback of the leader robot based on the follower robot side. $F_L \in R^{3\times1}$ is the total force in the leader robot side.

III. System description

The bilateral teleoperation system enhanced by the stable VG method is shown in Fig. 1. The proposed framework includes multimodal feedback-based bilateral teleoperation control, virtual guidance and robot-environment interaction. The human teleoperates the robot through the haptic device, and the exerted force by the human is f_h. The haptic device is controlled by the human operator, force feedback from the follower robot and virtual guidance, which could provide a cue for the human operator. Although the VG provides guidance force for human operator, the final command for haptic is determined by the human operator. Human operator could adjust the pose of the haptic device based on the virtual feedback and the desired trajectory. When the robot interacts with environments, the contact force f_e will be rendered by force feedback. The VG force is generated based on the virtual guidance model and current robot states, position and velocity. The VG model is attained through NEUM and a dynamic system.

A. Hardware system

As shown in 2, Franka robot is used to conduct the experiment, which is running at 1K Hz. The haptic Touch X is used as a teleoperation device, which runs at 1K Hz and provides haptic feedback for human operators. The RealSense camera (435i) is used to provide visual feedback for human operators. The joint torque sensors are used to calculate the interaction force between the robot end effector and the environment, running at 1k Hz. One laptop computer is used to control the robot, and another laptop is used to visualize the image feedback. For the ultrasound scanning experiment, an ultrasound probe is used for ultrasound scanning and an artificial Phantom is used to simulate the human body. One polishing tool is used to polish the component with an unknown surface, which is manufactured by 3D printing.

B. Software system

The bilateral teleoperation system consists of robot control, virtual guidance algorithm, bilateral teleoperation, etc. A laptop running a Linux system was used to implement the control algorithms. To integrate these different components, ROS Medolic is used to communicate among different modules: robot control, virtual guidance module, force feedback, and visual feedback. Libfrank and C++ are used to control the Franka robot at 1K Hz. The low-level controller of the Franka robot is implemented by the impedance controller in Cartesian space. The impedance controller can achieve compliant interaction between the end-effector of the robot arm and the objects. For the NEUM implementation details, please refer to our previous work [23].

IV. Experiment validation

We adopt two typical tasks, robot-assisted ultrasound scanning and robot-assisted polishing, to evaluate the performance of virtual guidance for bilateral teleoperation. In our previous work, we developed a robot-assisted ultrasound system and conducted a user study to compare the performance of various teleoperation interfaces [3]. The study revealed that the ultrasound scanning task is quite challenging for human operators. Therefore, in this work, we adopt this task as an experimental case. So in this work, we adopt this task as an experiment case. We present the details and the challenges of these tasks by pure bilateral teleoperation as follows.

A. Task description

a) Ultrasound scanning approaching, as shown in Fig. 3. In our setup of robot-assisted ultrasound scanning by teleoperation, we only use the monocular camera for visual feedback and the haptic device for force feedback. We conducted experiments on the Phantom. The robot from a random pose at the beginning achieves the desired position on Phantom, as shown in Fig. 3. In this task, the teleoperation needs to adjust the position and orientation simultaneously.

Fig. 3. The ultrasound scanning task (a) and polish task (b) by the bilateral teleoperation system.

b) Polishing task by the collaborative robot in an unknown environment continuously, as shown in Fig. 3. In our previous work, we studied impedance learning for robot-human-environment physical interaction [26]. The robot-assisted polishing task presents significant challenges for teleoperation demonstrations. In this work, we further evaluate the performance of the virtual guidance by a manufacturing application. In this task, the human operator needs to adjust the position and contact force simultaneously based on the visual and force feedback. The interaction with unknown interfaces and bilateral teleoperation without virtual guidance is still challenging for human operators nowadays.

V. Experiment study cases

A bilateral teleoperation system is developed, and the experiment setup can be seen in Fig. 2. Based on this experimental platform, we conduct two tasks: ultrasound scanning and robot-assisted polishing.

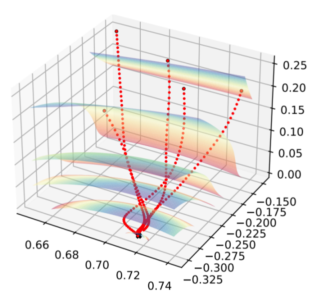

Fig. 4. Energy field for ultrasound scanning.

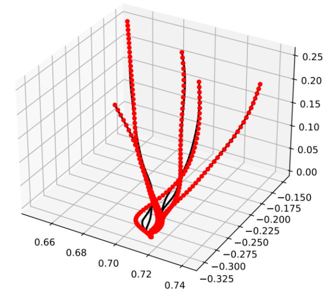

Fig. 5. Trajectories reproducing by energy fields for ultrasound scanning approach.

A. Medical examination - ultrasound scanning

As shown in Fig. 4, the energy function for ultrasound scanning is built based on human demonstration. The curved surfaces represent the energy fields, and the points on the same surface have the same energy. The red dotted lines are the demonstration trajectory for the building of virtual guidance, in which the data set has five demonstration trajectories. The black point is the target goal. As shown in

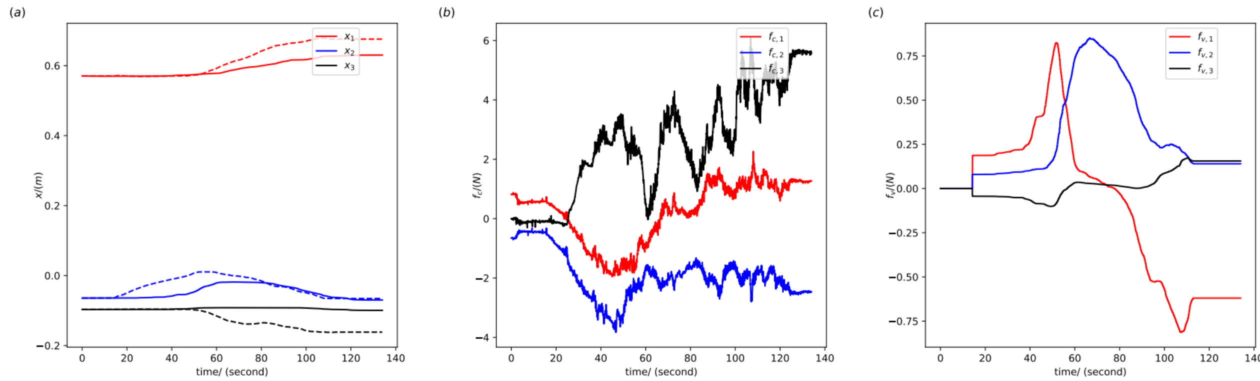

Fig. 6. Results of the polishing experiment under virtual guidance.

Fig. 5, the black lines are the reproducing trajectories, based on the energy field and the given target goal.

The method is also effective for the contact-rich task. As shown in Fig. 7, virtual guidance based on the energy fields is generated based on human demonstration. As shown in Fig. 8, only five demonstrations are required for the energy field and virtual guidance. The reproducing trajectories, black lines, demonstrate that the energy fields could generate virtual guidance to converge to the desired target.

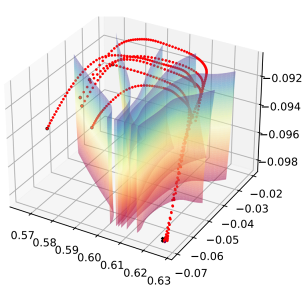

Fig. 7. Energy field for polishing.

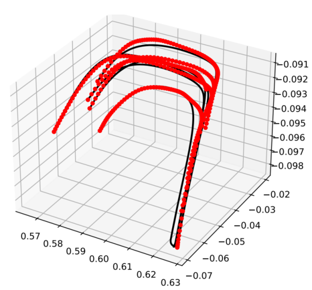

Fig. 8. Trajectories reproducing based on the energy field for polishing.

The experiment results of robot-assisted polishing are shown in Fig. 6. (a) shows the desired trajectories of the human operator and the actual trajectories of the polishing tool in x1 (X-axis), x2 (Y-axis), and x3 (Z-axis), respectively. (b) shows the actual contact forces of $f_{c,1}$, $f_{c,2}$,$f_{c,3}$ in X-axis, Y-axis, and Z-axis, respectively. (c) is the virtual guidance force,$f_{v,1}$, $f_{v,2}$,$f_{v,3}$.

The tracking error in X is around 4cm; when the human operator deviates from the desired position, the virtual guidance force will be large, as shown in (c). The tracking accuracy along the Z-axis is large because the polishing tools interact with the rigid environment 3D printing component. The contact force in the Z-axis is large, almost to 6N. However, the virtual guidance force is small, because the human operator exerts contact force along the Z-axis by deviating from the current position. Because the low-level controller is impedance control, the contact force can be generated through control of the position from the equilibrium point.

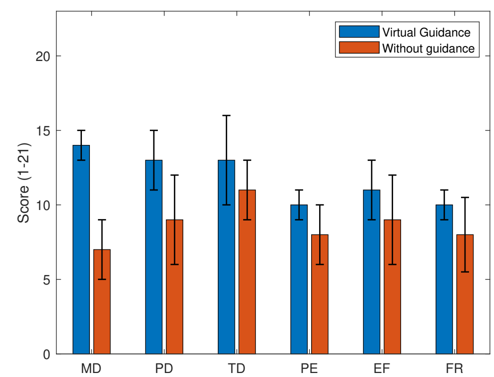

Fig. 9. Subjective evaluation based on NASA-TLX scores for teleoperation with virtual guidance and without.

B. Performance evaluation based on NASA-TLX

To compare the effectiveness of virtual guidance, user study experiments were conducted for the aforementioned tasks. During the experiment, participants were tasked with completing the designated tasks using direct teleoperation without guidance, as well as with virtual guidance. Following each task completion, participants were asked to fill out the NASA-TLX questionnaire, rating their experience across six aspects: mental demand (MD), physical demand (PD), temporal demand (TD), performance (PE), effort (EF), and frustration (FR). The scores are presented in Fig. 9 as a box plot, to graphically demonstrate the maximum, upper quartile, median value, lower quartile, and minimum value (shown as horizontal lines from top to bottom respectively) as well as mean value (shown as a cross point). Overall, users of direct teleoperation without guidance tended to assign lower scores across all aspects. The most significant contrast in scores is observed notably in mental demand (MD), where the mean value without guidance (13.4) nearly doubles that with virtual guidance (6.8). Similarly, discernible distinctions between virtual guidance and no guidance are apparent in physical demand (PD) and effort (EF), with differences of 4.2 and 4, respectively. Participants utilizing direct teleoperation without guidance also exhibited lower scores in performance (PE) and frustration (FR), suggesting that increased task demands and effort correlated with slightly diminished performance when compared to virtual guidance.

VI. Discussion and conclusion

In this paper, we proposed a stable virtual guidance method for teleoperation-based robot learning from demonstration. Two tasks, ultrasound scanning and robot-assisted polishing, have been conducted to evaluate the performance of the proposed method. The VG is designed by a neural energy function with a unique minimum, which can generate guidance force, from the high energy position to the low energy position for the teleoperation operator. The unique minimum feature can guarantee to converge to the desired position. In addition, the virtual guidance is built based on a human demonstration data set. We developed a bilateral teleoperation system for conducting user studies to evaluate its effectiveness.

Under VG guidance, human operators can mentally and physically reduce the workload. However, as mentioned above, it is hard to demonstrate one optimal trajectory through a one-shot demonstration based on visual and force feedback teleoperation, especially for the contact-rich tasks and dexterous manipulation tasks, which often require the operator to control the orientation, position and contact force simultaneously, such as polishing in the manufacturing and ultrasound scanning in robot-assisted medical examination. However, human operators can correct and interact through the human-in-the-loop mechanism to update the previous demonstration. In the future, this correction and human demonstration will be investigated to optimize the VG. In this case, the iterative optimization method can be used to update the NEUM, which will further provide better virtual guidance for human operators. We also plan to investigate the adaptive design of virtual guidance for different tasks to improve user experience.

VII. SUPPLEMENTAL MATERIALS

Supplementary Video S1

References

[1] W. Si, N. Wang, and C. Yang, "A review on manipulation skill acquisition through teleoperation-based learning from demonstration," *Cognitive Computation and Systems*, vol. 3, no. 1, pp. 1–16, 2021.

[2] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, 2006.

[3] W. Si, N. Wang, and C. Yang, "Design and quantitative assessment of teleoperation-based human–robot collaboration method for robot-assisted sonography," *IEEE Transactions on Automation Science and Engineering*, 2024.

[4] W. Si, Y. Guan, and N. Wang, "Adaptive compliant skill learning for contact-rich manipulation with human in the loop," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 5834–5841, 2022.

[5] J. Luo, Z. Lin, Y. Li, and C. Yang, "A teleoperation framework for mobile robots based on shared control," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 377–384, 2019.

[6] H. Hu, D. Shi, C. Yang, W. Si, and Q. Li, "A novel shared control framework based on imitation learning," in *2024 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2024, pp. 1–6.

[7] S. S. Restrepo, G. Raiola, P. Chevalier, X. Lamy, and D. Sidobre, "Iterative virtual guides programming for human-robot comanipulation," in *2017 IEEE International conference on advanced intelligent mechatronics (AIM)*. IEEE, 2017, pp. 219–226.

[8] P. Stotko, S. Krumpen, 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 *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 3630–3637.

[9] P. Chang, R. Luo, M. Dorostian, and T. Padır, "A shared control method for collaborative human-robot plug task," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7429–7436, 2021.

[10] W. Si, T. Zhong, N. Wang, and C. Yang, "A multimodal teleoperation interface for human-robot collaboration," in *2023 IEEE International Conference on Mechatronics (ICM)*, 2023, pp. 1–6.

[11] S. A. Bowyer, B. L. Davies, and F. R. y Baena, "Active constraints/virtual fixtures: A survey," *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 138–157, 2013.

[12] M. Selvaggio, G. Notomista, F. Chen, B. Gao, F. Trapani, and D. Caldwell, "Enhancing bilateral teleoperation using camera-based online virtual fixtures generation," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 1483–1488.

[13] G. Raiola, S. S. Restrepo, P. Chevalier, P. Rodriguez-Ayerbe, X. Lamy, S. Tliba, and F. Stulp, "Co-manipulation with a library of virtual guiding fixtures," *Autonomous Robots*, vol. 42, pp. 1037–1051, 2018.

[14] V. Pruks and J.-H. Ryu, "Method for generating real-time interactive virtual fixture for shared teleoperation in unknown environments," *The International Journal of Robotics Research*, vol. 41, no. 9-10, pp. 925–951, 2022.

[15] 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.

[16] F. Ferraguti, N. Preda, M. Bonfe, and C. Secchi, "Bilateral teleoperation of a dual arms surgical robot with passive virtual fixtures generation," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 4223–4228.

[17] C. J. Pérez-del Pulgar, J. Smisek, V. F. Munoz, and A. Schiele, "Using learning from demonstration to generate real-time guidance for haptic shared control," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2016, pp. 003 205–003 210.

[18] Z. Lu, W. Si, N. Wang, and C. Yang, "Dynamic movement primitives-based human action prediction and shared control for bilateral robot teleoperation," *IEEE Transactions on Industrial Electronics*, 2024.

[19] J. Dong, W. Si, and C. Yang, "A novel human-robot skill transfer method for contact-rich manipulation task," *Robotic Intelligence and Automation*, vol. 43, no. 3, pp. 327–337, 2023.

[20] W. Si, C. Guo, N. Wang, M. Yang, R. Harris, and C. Yang, "A unified deep imitation learning and control framework for robot-assisted sonography," in *2023 IEEE International Conference on Development and Learning (ICDL)*, 2023, pp. 318–323.

[21] D. Papageorgiou, F. Dimeas, T. Kastritsi, and Z. Doulgeri, "Kinesthetic guidance utilizing dmp synchronization and assistive virtual fixtures for progressive automation," *Robotica*, vol. 38, no. 10, pp. 1824–1841, 2020.

[22] J. Luo, W. Liu, W. Qi, J. Hu, J. Chen, and C. Yang, "A vision-based virtual fixture with robot learning for teleoperation," *Robotics and Autonomous Systems*, vol. 164, p. 104414, 2023.

[23] Z. Jin, W. Si, A. Liu, W.-a. Zhang, L. Yu, and C. Yang, "Learning a flexible neural energy function with a unique minimum for globally stable and accurate demonstration learning," *IEEE Transactions on Robotics*, 2023.

[24] K. I. Penev, *A potential field method for the automatic design of modular fixtures*. University of Southern California, 1997.

[25] Y. Michel, R. Rahal, C. Pacchierotti, P. R. Giordano, and D. Lee, "Bilateral teleoperation with adaptive impedance control for contact tasks," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5429–5436, 2021.

[26] X. Xing, E. Burdet, W. Si, C. Yang, and Y. Li, "Impedance learning for human-guided robots in contact with unknown environments," *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3705–3721, 2023.