Adaptive Compliant Skill Learning for Contact-Rich Manipulation With Human in the Loop

SI, Weiyong, GUAN, Yuan <http://orcid.org/0000-0001-9040-3169> and WANG, Ning <http://orcid.org/0000-0002-3264-1852>

Adaptive Compliant Skill Learning for Contact-Rich Manipulation with Human in the Loop

Weiyong Si, *Student Member, IEEE,* Yuan Guan, *Student Member, IEEE,* and Ning Wang, *Member, IEEE*

***Abstract*—It is essential for the robot manipulator to adapt to unexpected events and dynamic environments while executing the physical contact-rich tasks. Although a range of methods have been investigated to enhance the adaptability and generalization capability of robot manipulation, it is still difficult to perform complex contact-rich tasks, e.g., rolling pizza dough and robot-assisted medical scanning, without the assistance from a human in the loop. We proposed a novel framework combining learning from demonstration (LfD) and human experience to enhance the safety and adaptability of the robot manipulation. In this framework, dynamic movement primitives (DMPs) is employed for manipulation skills learning from demonstrations, and human correction is applied to update the pre-trained DMPs skills model. We conducted experiments on the Franka Emika Panda Robot with pizza dough rolling tasks. The results demonstrate that the proposed framework could effectively improve the performance of the physical contact-rich tasks, and the human correction method through teleoperation provides a potential solution for advanced interaction tasks with complex and dynamic physical properties.**

***Index Terms*—Bilateral teleoperation; Compliant movement skills; Learning from demonstration; Force and impedance control.**

I. INTRODUCTION

ROBOT manipulators have been employed in a range of fields, e.g., industry, medical examination and space exploration etc. However, robot manipulators are expected to deal with more complex contact-rich tasks; For example, robots continuously interact with the environment physically, which is ubiquitous during daily manipulation [1] [2] [3]. In addition, for the autonomous execution of robot manipulators, the programing manual is time-consuming and in-flexible for dynamic tasks or environments. Recently, as data-driven techniques, e.g., deep learning, imitation learning and reinforcement learning, are successfully employed in the robotic community [4], the machine learning methods can also be used to transfer human-like skills to robot manipulators [5] [6].

The contact-rich tasks are characterized by high variability in dynamic tasks or environments, making it challenging to interact autonomously [1]. Nowadays, the reactive mechanism has been integrated with online planning and control to deal with these tasks [7]. For example, optimization-based reactive planning was proposed to generate the real-time trajectories, and then the torque-based controller was designed to track the generated trajectories [8]. Also, data-driven methods were used to process online perceptual feedback, especially deep learning technology, which is employed to fuse multimodal feedback. However, the real-time optimization-based methods often require high computation capacity for the multiple degrees of freedom (DoF) manipulators. In addition, optimization-based methods usually assume that the interaction model is known. Nevertheless, it is hard to attain an accurate interaction model due to the complex physical properties of the interaction process. Reinforcement learning (RL), especially deep RL (DRL), has been proposed to deal with complex interaction tasks [9]. Deep learning has been successfully used for machine vision, perceptual information extraction, etc. The integration of deep learning and RL can further improve the generalisation capability of multimodal perception information [10]. However, the DRL often requires a large amount of data, which is difficult and costly to attain for robot in practice. Hence, it remains difficult for robot manipulators to perform contact-rich tasks autonomously [11].

Teleoperation has been widely used in robotics [12], especially when dealing with tasks in hazardous environments or remote medical diagnostics, such as ultrasound scanning

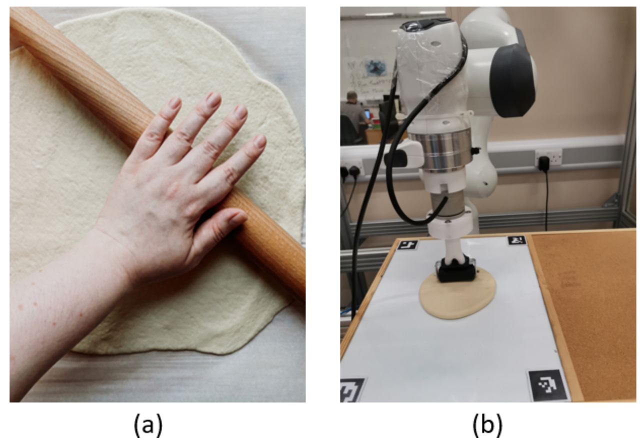


Fig. 1: Rolling the pizza dough by a robot. The dough often varies in hardness, requiring the roller to interact with different forces. When humans perform this task, we can naturally adjust force, impedance, and motion. When the robot interacts with the environment, which is deformable and difficult to model, the robot also needs to adjust the force, impedance and motion simultaneously.

Manuscript received: October 6, 2021; Revised January 14, 2022; Accepted February 19, 2022.

W. Si, Y. Guan and N. Wang are with the Faculty of Environment and Technology and Bristol Robotics Lab at the University of the West of England, Bristol, BS16 1QY, UK. (Corresponding author is Ning Wang. Email: Katie.Wang@brl.ac.uk)

This paper was recommended for publication by Editor Jens Kober upon evaluation of the Associate Editor and Reviewers' comments.

Digital Object Identifier (DOI): see top of this page.

[13]. Recently, one flexible solution has been integrating humans with superior experience and cognitive abilities into the execution loop of robots through teleoperation [14]. In addition, robot skill learning from demonstration via teleoperation enables the human-robot skill transfer manipulative in a remotely feasible manner, and the human-in-the-loop mechanism can also improve the safety of the robot execution, especially for the physical contact tasks. However, unlike the demonstration by observation or kinesthetic demonstration [15], the learning from demonstration by teleoperation needs the extra input devices. In terms of the teleoperation interface for human-robot skill transfer, immersive teleoperation interfaces have been developed to improve the intuitiveness of the human demonstration [16] [17]. Benefiting from various interactive devices, such as the Touch from 3D Systems, force feedback information makes it possible for human operators to perceive the process of robot-environment interaction [18]. The haptic devices providing tactile feedback also improve the transparency of bilateral teleoperation [19]. And, this advantage could increase the safety of teleoperation. However, the teleoperation device also causes problems such as time delays and significant cognition workloads, especially configuration differences between leader and follower devices [20].

To address the abovementioned problems in bilateral teleoperation, several researchers have investigated interface design in terms of software as well as hardware to reduce the cognitive workload [21]. An intuitive teleoperation interface is the key to transfer human-like skills from humans to robots [22]. For example, the authors introduced a subspace method in teleoperation, which improves the accuracy of teleoperation and reduces the cognitive workload [23] [24]. However, most of the existing work on skill learning through teleoperation has focused on learning kinematic skills, such as motion skill and velocity skill, etc [25]. It is still difficult to transfer compliant skills from humans to robots.

In this paper, we present a framework for transferring compliant manipulation skills from humans to robots through bilateral teleoperation and investigate the human-in-the-loop mechanism for robot manipulation skill learning and adaptation. First, we develop a teleoperation system based on the collaborative robot, haptic device, depth camera, and force/torque sensor for human-robot skill transfer through teleoperation. The multimodal feedback, interaction force, and visual feedback are employed to enhance the transparency of bilateral teleoperation. In addition, the human-in-the-loop mechanism provides a solution to demonstrate and correct the behaviour of autonomous robots. Compared to the literature [7], the human-guided correction method can exploit human decision-making and cognitive abilities to improve the adaptability and generalization of the skill model. The main contributions of this paper can be summarized as follows.

- We developed a hybrid control architecture for a bilateral teleoperation system that includes hybrid force and position control as well as impedance control to achieve human-robot compliant skill transfer.
- The teleoperation-based system allows human operators to correct the behaviour of the autonomous robot. The updated behaviour is employed to update the pre-learned compliant skills model.
- The proposed solution is evaluated on the pizza dough rolling of different hardness. The updated skill model can increase the performance of the dough rolling, such as the success rate, and uniform force on the dough. This solution also has the potential to extend to other contact-rich tasks, such as medical scanning.

The remaining of the paper is as follows: Section II presents the preliminary knowledge of robotic dynamics and teleoperation control. The proposed framework is developed in Section III. The design of the experiment and the results and performance analysis are present in Section IV. Finally, discussions and future work are given in Section V.

II. PRELIMINARY

A. Robotic dynamics and control

The dynamics of the general serial manipulator robot in Cartesian space can be modelled as [26],

$$M(q)\ddot{x} + C(q,\dot{q})\dot{x} + G(q) = f_c + f_{ext} \tag{1}$$

where the $M(q)$ is the inertia matrix in Cartesian space, the $C(q,\dot{q})$ is the Coriolis term and $G(q)$ represents gravitational force. f_c is the control force and f_{ext} is the interaction force with the environment. q and $\dot{q}$ represent the joint position and velocity, respectively, the $\dot{x}$ and $\ddot{x}$ are the velocity and acceleration of robot end-effector in Cartesian space.

For the bilateral teleoperation system, the dynamics of the leader robot can be described as [27],

$$M_m(q_m)\ddot{x}_m + C_m(q_m,\dot{q}_m)\dot{x}_m + G_m(q_m) = f_m + f_h \tag{2}$$

where $M_m(q_m)$ is the inertia matrix of leader robot, $C_m(q_m,\dot{q}_m)$ is the Coriolis and centrifugal terms and $G_m(q_m)$ represents the gravitational force. f_m and f_h are the control and operator force, respectively. q_m and $\dot{q}_m$ are the joint position and velocity in the joint space, respectively, the $\dot{x}_m$ and $\ddot{x}_m$ represents the velocity and acceleration of end-effector in Cartesian space. Similarly, on the remote manipulator side, the dynamics of the follower robot can be described as,

$$M_r(q_r)\ddot{x}_r + C_r(q_r,\dot{q}_r)\dot{x}_r + G_r(q_r) = f_r + f_e \tag{3}$$

where the $M_r(q_r)$ is the inertia matrix of the follower robot, $C_r(q_r,\dot{q}_r)$ is the Coriolis and centrifugal term and $G_r(q_r)$ represents the gravitational force. f_r and f_e are the control and interaction force executed on the robot, respectively. q_r and $\dot{q}_r$ are the joint position and velocity in the joint space, respectively, the $\dot{x}_r$ and $\ddot{x}_r$ represents the velocity and acceleration of end-effector in Cartesian space. For both the leader and follower robots, $(\dot{M}_i - 2C_i)$ for $i = \{r, m\}$ are skew symmetry, which represents that the remote manipulator and the leader device are passive respectively. The control command for the remote manipulator is generated by an impedance controller,

$$f_r = K_r(x_m - x_r) - D_r\dot{x}_r \tag{4}$$

where K_r is the stiffness matrix, the D_r represents the damping matrix. x_m and x_r are the position of the leader and follower robots, respectively. For the leader robot, the force

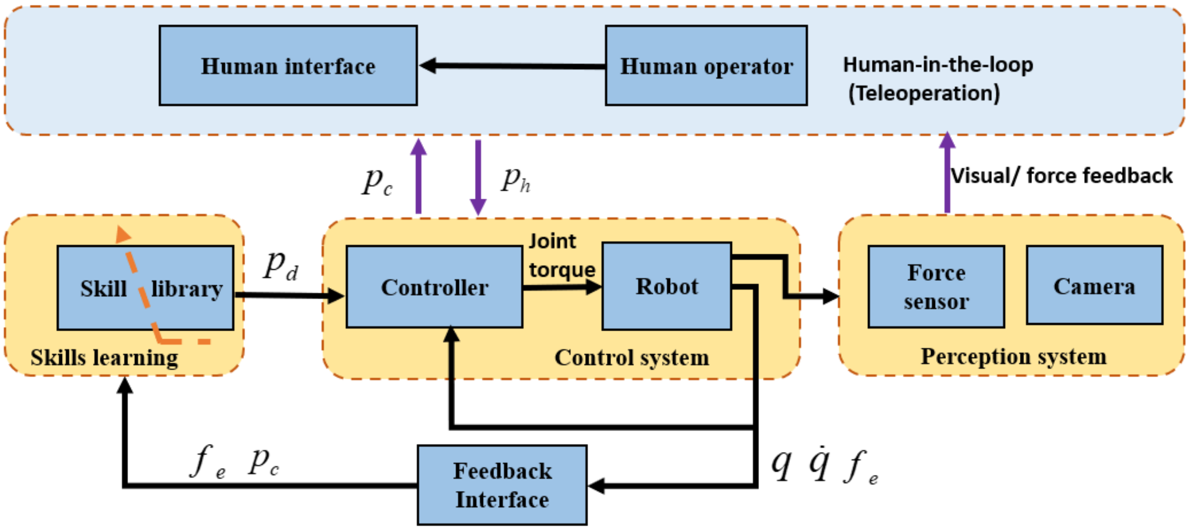


Fig. 2: The overview of the proposed framework. The visual and force feedback provided perceptual information to the human operator. The control system consists of impedance control and hybrid force and position control to generate the joint torque. More details of the control system can be found in Fig.3. The skills learning module encodes the manipulation profile f_e, p_c in Cartesian space and generates desired pose p_d. The teleoperation module monitors the process and provides human correction p_h when necessary.

feedback is designed to reflect the interaction force between the follower robot and environment,

$$f_m = K_m(-f_e - D_m \dot{x}_m) \tag{5}$$

where D_m is the damping matrix. K_m is the scaling parameter. The stability of the bilateral teleoperation system can be proved by passivity analysis [27]. The bilateral teleoperation could provide force feedback for the human operator, which could allow the adaptive impedance skill transfer between the human operator and robot.

III. METHODOLOGY

A. Dynamic movement primitive

Dynamic movement primitives (DMPs) was proposed by Ijspeert to study motor control of humans, which was inspired by the dynamic systems and human motor control. The essential of DMPs was to encode the manipulation skills by a dynamic system. Since then, a number of improved DMPs have been proposed, such as orientation DMPs [28] and constrained DMPs [29] [30]. In addition, merging the separate DMPs into a complex manipulation was also investigated [28]. DMPs can also be employed to model multiple DoFs system, each DoF can be modelled separately, and a canonical system achieves the coupling among these DoFs. For readability, we introduce the DMPs for one degree of multiple dynamic systems,

$$\begin{aligned} \tau_s \dot{v} &= \alpha_z(\beta_z(p_g - p) - v) + F(x) \\ \tau_s \dot{p} &= v \end{aligned} \tag{6}$$

where the p_g represents the goal position, p represents the current position; τ_s is the scaling parameter, the v is the velocity, α_z, β_z need to be designed, however, these parameters satisfying $\alpha_z = 4\beta_z$. $F(x)$ is the nonlinear forcing term, which is used to affect the trajectory. Generally, the $F(x)$ can be be composed of the Gaussian functions,

$$F(x) = \frac{\sum_{i=1}^{N} \psi_i(x) w_i}{\sum_{i=1}^{N} \psi_i(x)} x(p_g - p_0) \tag{7}$$

$$\psi_i(x) = \exp(-h_i(x - c_i)^2) \tag{8}$$

A canonical system was used to determine the phase variable x, and the canonical system can coordinate different DoFs, which can be described as,

$$\tau_s \dot{x} = -\alpha_x x, \quad x \in [0,1]; \quad x(0) = 1 \tag{9}$$

where τ_s represents the scaling parameter, α_x is a positive coefficient, and the initial value of x is $x(0) = 1$, which can converge to zero exponentially.

The original DMPs model is used to encode the manipulation skill offline, and it is not suitable for online skill update. Recently, a sensory-based coupling term was proposed to achieve reactive planning and control. In addition, the coupling term could be used to avoid obstacles based on the real-time perception feedback. Inspired by the work [7], we proposed a novel formation of DMPs for robot skill update online through human-in-the-loop,

$$\begin{aligned} \tau_s \dot{v} &= \alpha_z(\beta_z(p_g - p) - v) + F(x) + H(\cdot) \\ \tau_s \dot{p} &= v \end{aligned} \tag{10}$$

where $H(\cdot)$ is the human interaction terms, which provides a human interface to update the pre-defined skill online. This human interaction term is activated only in the shared control mode when humans interact with the execution process where the pre-trained skill cannot generalise to new cases. The $H(\cdot)$ is approximated by a set of radial basis functions as same as $F(x)$, and the locally weighted projection regression (LWPR) technique is used to calculate the weights of this term $H(\cdot)$.

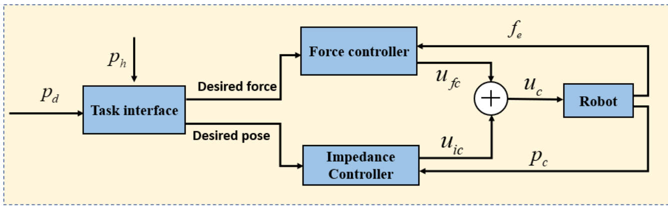


Fig. 3: The architecture of the control system. The required poses and forces are derived from the skill model and teleoperation. The task interface decouples the tasks to each DoF, and the desired commands are fed to the corresponding controllers. The computed torques from the force and impedance controllers are combined and provided to the robot.

Moreover, this term is modified only in the specific DoF, and this interaction term could be updated iteratively.

$$H(\cdot) = \begin{cases} h(s) & \gamma = 1 \\ 0 & \gamma = 0 \end{cases} \tag{11}$$

where $h(s)$ is the output of interaction term trained based on the human correction. $\gamma = 0$ and $\gamma = 1$ represent autonomous mode and human correction mode, respectively.

For the multi-step tasks, it often needs to merge several DMPs. In this work, we adopt the first method described in [28], which could guarantee position and orientation to be smooth. This method requires fewer parameters than other methods, e.g., smooth acceleration transition, and it is feasible for multi-step merging without high requirement on efficiency.

$$p_{ne} = p_{pr} \qquad v_{ne} = v_{pr} \tag{12}$$

$$q_{ne} = q_{pr} \qquad \omega_{ne} = \omega_{pr} \tag{13}$$

where the p_{ne}, v_{ne}, q_{ne}, ω_{ne} are the initial position, velocity, angular and angular rate of next DMP respectively; the p_{pr}, v_{pr}, q_{pr}, ω_{pr} are for the last DMP. For the orientation control, we adopt the unit quaternion based DMPs [28] to model the orientation skills. The orientation generalisation was not considered in this work; the details of orientation DMPs can refer to [28].

B. Hybrid force and position control

Impedance control, proposed by Hogan in 1985 [31], has been investigated widely for robot manipulator control, such as compliant control, human-robot interaction, etc. The impedance controller is similar to a sping-damper dynamic system, which is used to model the dynamic interaction between the robot and its environment. This control of dynamic behaviour allows human-robot and robot-environment safety interaction. In addition, adaptive impedance control has proven an effective approach to improve control performance during the interaction between robots and the environment [32] [27] [2]. The impedance controller can be implemented in the joint space and Cartesian space. For the tasks of end-effector interaction with the environment, the impedance controller in Cartesian space is more suitable. We implemented the impedance control in the task space, and the control law can be described as,

$$u_{ic} = -J^T(q)(K\Delta x_t + D\Delta\dot{x}_t + k_t) \tag{14}$$

where the u_{ic} is the command generated by the impedance controller, the $J(q)$ is the Jacobian matrix associated with robot configuration, K and D are the stiffness and damping matrix, which is used to determine the characteristic of impedance. Δx is the position error between the current position and the desired position; $\Delta\dot{x}$ is the velocity error between the current velocity and the desired velocity. k_t is the feedforward compensation calculated by the dynamic model of manipulator.

The hybrid force-position approach investigates the two reciprocal subspaces: twists and wrenches, regulating the contact force and simultaneously tracking the desired motion. For the force controller part, a proportional-integral method can be stable for a high-stiffness control process. The force controller can be described as [33],

$$\Delta f = \Im_{ee}(t) - \Im_d(t) \tag{15}$$

$$u_{fc} = J^T(q)[K_{fp}\Delta f + K_{fi}\int \Delta f dt] \tag{16}$$

where the $\Im_{ee} = (f_{ee}^T, m_{ee}^T)$ is the wrench measured in the end-effector coordinate system, generated by interacting with the environment. The f_{ee} and m_{ee} are the current force and torque measured from torque/force sensor in this work. $\Im_d = (f_d^T, m_d^T)$ is the desired wrench, f_d and m_d are the desired force and torque. The command to the robot is the combination of force controller and the impedance controller. The combined command u_c of the robot to execute a given task can be,

$$u_c = u_{ic} + u_{fc} \tag{17}$$

where the u_{ic} is the command generated by the impedance controller, and the u_{fc} is the command generated by the force controller.

Based on the hybrid force-position controller, the control allocation for the motion tracking and force control is realized by a task matrix. We defined the task matrix as the control allocation matrix for the n-DOF to realize the switch among different tasks. We introduce the position control matrix M_p; hence the force control matrix M_f can be described as,

$$M_f = I_6 - M_p \tag{18}$$

$$M_p = diag(s_i) \tag{19}$$

where the task matrix, M_f and M_p, are diagonal matrix, with the $s_i \in \{0, 1\}$, $s_i = 1$ represents position control in this direction, and $s_i = 0$ represents force control in this direction.

Different controllers are usually used for a multi-step task for different stages, for example, impedance control in the approaching stage and hybrid force and position control in the rolling stage. It is important to design the switching transition strategy when switching between different controllers to ensure smooth transitions and reduce jitter. The transition of two controllers, u_1 and u_2, is realized through the linear interpolation during the transition windows,

$$u = (1-\beta)u_1 + \beta u_2 \tag{20}$$

where u is the control command during the transition. $\beta{=}t/T$ and $t \in [0, T]$. The linear interpolation allows for a smooth transition between two different controllers.

C. Human correction interface

In [14], the authors proposed a shared control interface to mix human input and autonomous commands. The modification strategy by human operators can be described as,

$$x = x_n + \delta y \tag{21}$$

where $x_n \in R^m$ is the nominal robot state, and $\delta y \in S(R^m)$ is the modification to the robot state variable. $S(R^m)$ is the correction interface, the input is the modification command, and the output of the correction interface is the modified task variables. A nominal task model can be learned through demonstration offline. A corrective command is generated by the human operator based on the observations of any errors caused by the robot state or the environment state.

The human-in-the-loop allows humans to intervene autonomous execution of the robot through the teleoperation input device. The human correction command can be modelled an ordinary differential equation about the input of teleoperation device [14],

$$\ddot{\delta}y + b_c\dot{\delta}y + k_c\delta y = u \tag{22}$$

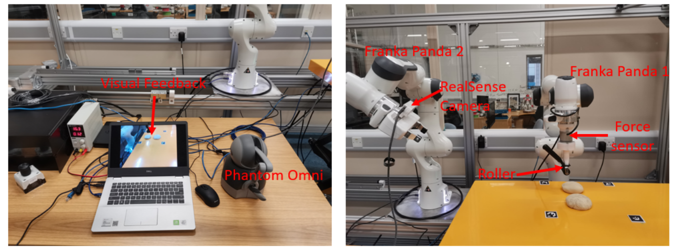
where δy represents the human command, k_c and b_c are the stiffness and damping parameter of the human correction dynamics, u represents the output of the haptic device.

The force/torque sensor has noise; therefore, it is necessary to filter the measurement data. In this work, we adopted the Kalman filter to reduce the noise, and the Kalman filter's update equations can refer to [34]. Other robot state variables, e.g., position and velocity, are used directly from the data provided by the Franka Control Interface (FCI), without additional filtering processing.

IV. EXPERIMENT STUDY CASES

We evaluated the proposed framework through a typical task in life, rolling pizza dough. This task consists of the interaction of the robot with its environment, a task that requires both appropriate contact force control and trajectory control simultaneously. In addition, it requires the robot to adapt its behaviour to the shape and hardness of the dough. However, the exact hardness of the dough is practically difficult for robots to attain in advance. The experimental setup is shown in Fig.4, and the main components of the experimental platform are as follows.

- A 7-DoF Franka Emika Panda 1 is used to carry a roller to perform the dough rolling task.
- A haptic device, Touch from 3D Systems, is used as the input device of teleoperation. The Touch has 6 DoFs and two extra state control buttons. The button states are used to switch among the control modes. The pose of Touch is used to correct the robot state in shared control mode.
- Realsense camera (D435i), carried by Franka Emika Panda 2, captures remote work scenes. The camera could provide visual feedback for human teleoperation.
- Control computer running Ubuntu 18.04, which is connected to the Touch device, force/torque sensor, Franka Emika Panda 1, and the camera. The ROS Melodic is used to communicate with different devices.
- The force/torque sensor is used to measure the interaction force during the tasks.



Leader site **Robot site**

Fig. 4: The setup of experimental platform.

(a) (b)

Fig. 5: (a) The robot rolls the pizza dough with proper force and motion. (b) The robot interacts with the soft dough with a large force, so the roller gets stuck. In this case, the forces along the X and Y axes are too large, which means that the task has failed.

A. Rolling dough skill transfer through bilateral teleoperation

In terms of the demonstration through teleoperation, we decoupled into two steps, motion skills in the X-Y plane and force skills along the Z-axis. First, we teleoperate the robot to roll the dough through the Touch input device. The translation mapping between the Touch and the end-effector is direct mapping with a scaling parameter in Cartesian space. And the demonstrated trajectory of the end-effector in the X-Y plane is recorded. The demonstration trajectory is used to encode the nominal skills in X-Y plane motion by the DMPs model. The learned skill can generate motion trajectories in the X-Y plane for the robot; hence the motion of the robot on the X-Y plane can be autonomous execution. The autonomous execution in the X-Y plane requires only the current coordinates and goal coordinates in the X-Y plane.

In the following demonstration, the Touch is used to demonstrate the force control along the Z-axis alone, while the

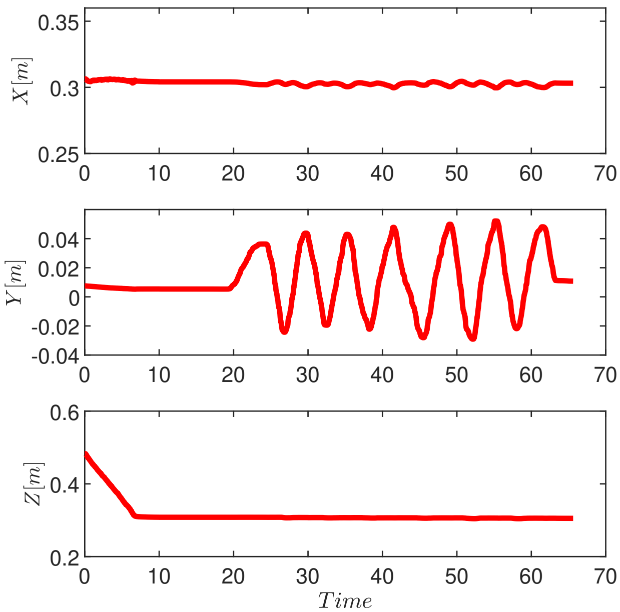


Fig. 6: The trajectories along the X, Y, Z axes in teleoperation demonstration.

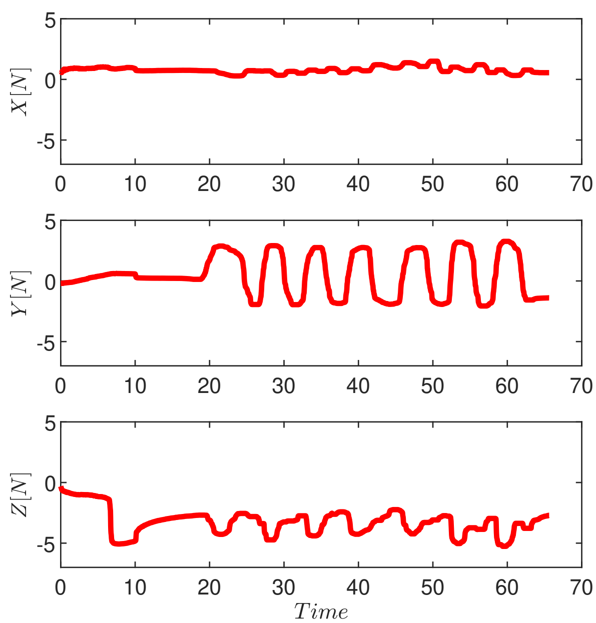


Fig. 7: The interaction forces along the X, Y and Z axes during the demonstration stage.

motion in the X-Y plane is generated by the learned DMPs. In this shared control mode, the human operator only teleoperates the contact force control, which could reduce the cognitive work and improve the accuracy of teleoperation. In addition, this also reduces communication traffic, which is helpful to reduce the time delay in the human-robot skill transfer, especially for the bilateral teleoperation in a long distance.

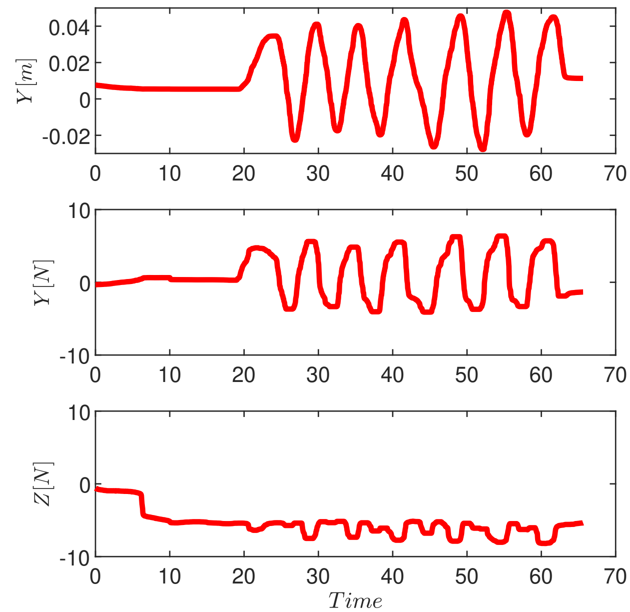


Fig. 8: The robot fails on the soft dough with learned force skill. The trajectory along Y-axis and the interaction forces along the Y-axis and Z-axis. The force along the Y-axis (direction of motion) is larger than 6N (normally, the force should be within 3N), which means that the roller is stuck.

The human could demonstrate to the robot the human-like manipulation skill, how to roll the dough only based on the force and visual feedback. To reproduce the learned skill, we let the robot roll the pizza dough with the same hardness autonomously. The interaction forces along the X, Y and Z-axis are shown in Fig.7.

B. Robot rolling soft doughes

When the robot rolls soft pizza dough, it needs to adjust the contact force according to the hardness of the dough. It cannot accomplish the task only relying on the learned force skill offline, for example, the large contact force causing the roller stuck, as shown in Fig. 8 and Fig.5. Although adaptive force control may deal with the uncertainty of hardness, it is hard to design a controller to deal with all kinds of hardness, especially since the properties of the dough are hard to model and attain in advance. In this case, the human could correct the contact force on top of the learned force skill through a shared control mechanism. For example, humans can reduce the contact force to generate an appropriate contact force for soft doughs. The human corrects the contact force based on the interaction force between the roller and the dough (force feedback via the Touch) and the deformation of the dough shape (visual feedback via the camera).

We evaluated the learned force skill on a novel dough that is softer than the one used in the demonstration phase to evaluate the generalization capability to the different hardness of the dough. As shown in Fig. 8, this is a failed case where the roller is stuck because the contact force is too large. Because

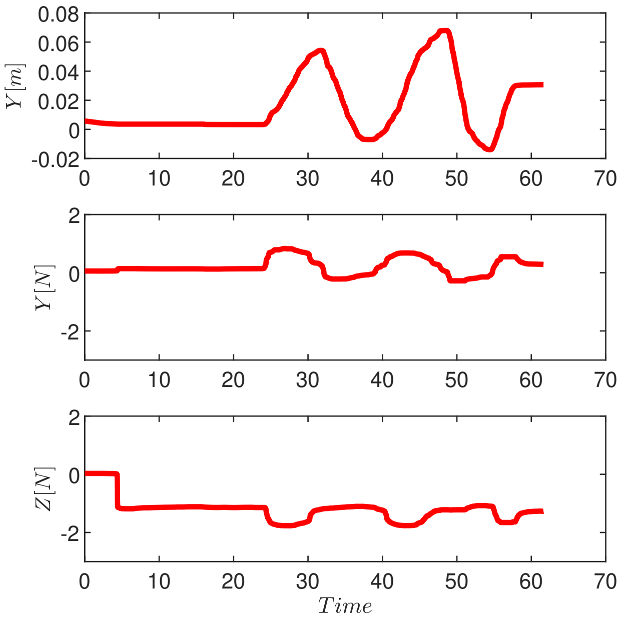


Fig. 9: The human corrects the contact force along the Z-axis through teleoperation online for the soft dough. The contact force along the Y-axis is less than 1 N, and the contact force along the Z-axis is less than the contact force for stiff dough (around 5N).

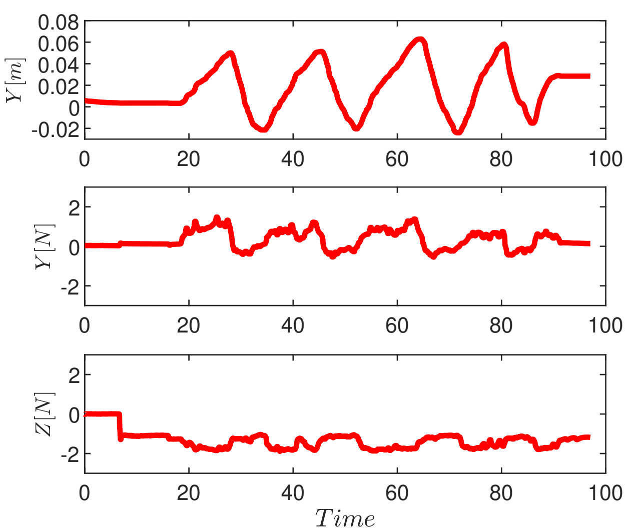


Fig. 10: Successs on the soft object.(a) is the trajectory in X,Y,Z. (b) is the interaction force. The contact force is reduced to roll on the soft dough.

the roller only moves along the Y-axis, we provide the motion along Y-axis and contact forces along Y-axis and Z-axis. The force along the Y-axis is too large, which means the task fails, and the roller is stuck. In addition, we tested on a novel dough, which is stiffer than the one used in the demonstration. However, the robot could not accomplish the task in the given time, and it needed more time to roll because the contact force was too small. In these cases, human correction is very useful to update the force skills. Compared to previous work, the human-guided method to update the skills model is more efficient. In addition, the human-guided method can also ensure safety.

For the above case, the human corrects the contact force on top of the autonomous commands (normal force) to generate an appropriate contact force for the soft dough. The contact force perceived by the force sensor is recorded to update the force skill model. The difference between the corrected contact force and the nominal force skill is used to train the human interaction term $H(\cdot)$ in the DMPs. As shown in Fig. 9, for the soft dough, the force along the Y-axis is less than 1N, and the corrected contact force along the Z-axis is around 2N. In this work, although we showed the human operator correct the contact force alone through the human-in-the-loop mechanism, this method can also be employed to update other variables, such as the orientation and motion, to meet some specific requirements in contact-rich applications.

We test the updated skill model with another soft dough. As shown in Fig. 10, the robot can roll the new soft dough autonomously. The contact force along the Y-axis is less than 2N, which means that the task is successful. Moreover, the motion pattern along the Y-axis is reproduced, which is similar to the learned pattern in the demonstration. Hence, the updated force skill can succeed in the novel doughs with soft hardness. In addition, the novel skills for soft dough can be put into the skill library as a new skill. This online correction and learning mechanism can expand the skill library built offline. This framework provides a solution to correct the robot's behaviour online through the developed interface, and the modified behaviour can be learned to update the skill library.

V. Discussion and Conclusion

In this paper, we developed a bilateral teleoperation system that includes a haptic device, collaborative manipulators, and force sensors etc. This system can be used to transfer contact-rich skills from humans to robots in a remotely feasible manner. And it can provide a human interaction interface for the human operator to interact and correct the robot's behaviour on top of the autonomous commands. Robots learning to manipulate deformable objects such as dough and medical scanning is difficult, especially since the interaction process is hard to model in advance. These tasks often include dynamic skills learning, such as force and stiffness, as well as kinematic skills. Therefore the force feedback is essential for these tasks. In addition, in some cases, humans cannot enter hazardous environments. The teleoperation-based human-robot skill transfer provides the solution to deal with this problem. We adopted rolling dough as a task to evaluate the performance of this solution, and the results show that the solution can deal with the uncertainty of the soft object.

In the beginning, the human could teach the robot to roll pizza dough through the shared control framework. To deal with the soft pizza dough, the human could correct the force command on top of the nominal skill output. And the correction profiles are used to update the pre-learned skill.

The updated skills are put into the skill library. Compared with work [14], we exploited the correction behaviour and provided the skill updated mechanism based on the improved DMPs model. In addition, unlike the DRL method, the human-guided skill learning and updating approach is more efficient. Especially in some cases, it is not feasible to try and update the skill model, such as medical scanning. The human-in-the-loop mechanism involving the human ensure safety, and the human interaction will benefit the robot skill learning.

Although this method can be used to correct the robot's behaviour online, it relies on the human to determine if human intervention is needed. Therefore, we will study the automated monitoring method to reason about the failure scenarios and then have a human start correcting the behaviour. Currently, this approach is used to modify the contact force, and we will also investigate how to correct other control variables, such as orientation and translation, simultaneously. The correction interface can be optimized to improve intuitiveness and flexibility.

References

[1] J. Cui and J. Trinkle, "Toward next-generation learned robot manipulation," *Science Robotics*, vol. 6, no. 54, p. eabd9461, 2021.

[2] Y. Li, G. Ganesh, N. Jarrassé, S. Haddadin, A. Albu-Schaeffer, and E. Burdet, "Force, impedance, and trajectory learning for contact tooling and haptic identification," *IEEE Transactions on Robotics*, vol. 34, no. 5, pp. 1170–1182, 2018.

[3] C. Yang, G. Peng, Y. Li, R. Cui, L. Cheng, and Z. Li, "Neural networks enhanced adaptive admittance control of optimized robot–environment interaction," *IEEE transactions on cybernetics*, vol. 49, no. 7, pp. 2568–2579, 2018.

[4] Z. Lu, N. Wang, and C. Yang, "A novel iterative identification based on the optimised topology for common state monitoring in wireless sensor networks," *International Journal of Systems Science*, vol. 53, pp. 1–15, 2021.

[5] O. Kroemer, S. Niekum, and G. Konidaris, "A review of robot learning for manipulation: Challenges, representations, and algorithms." *J. Mach. Learn. Res.*, vol. 22, pp. 30–1, 2021.

[6] C. Yang, C. Zeng, C. Fang, W. He, and Z. Li, "A dmps-based framework for robot learning and generalization of humanlike variable impedance skills," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 3, pp. 1193–1203, 2018.

[7] A. Rai, G. Sutanto, S. Schaal, and F. Meier, "Learning feedback terms for reactive planning and control," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 2184–2191.

[8] M. Vochten, W. Decré, E. Aertbeliën, and J. De Schutter, "Shape-preserving and reactive adaptation of robot end-effector trajectories," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 667–674, 2021.

[9] S. A. Khader, H. Yin, P. Falco, and D. Kragic, "Stability-guaranteed reinforcement learning for contact-rich manipulation," *IEEE Robotics and Automation Letters*, vol. 6, no. 1, pp. 1–8, 2020.

[10] M. A. Lee, Y. Zhu, P. Zachares, M. Tan, K. Srinivasan, S. Savarese, L. Fei-Fei, A. Garg, and J. Bohg, "Making sense of vision and touch: Learning multimodal representations for contact-rich tasks," *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 582–596, 2020.

[11] K. Chatzilygeroudis, V. Vassiliades, F. Stulp, S. Calinon, and J.-B. Mouret, "A survey on policy search algorithms for learning robot controllers in a handful of trials," *IEEE Transactions on Robotics*, vol. 36, no. 2, pp. 328–347, 2019.

[12] J. Luo, D. Huang, Y. Li, and C. Yang, "Trajectory online adaption based on human motion prediction for teleoperation," *IEEE Transactions on Automation Science and Engineering*, pp. 1–8, 2021.

[13] F. Abi-Farraj, T. Osa, N. P. J. Peters, G. Neumann, and P. R. Giordano, "A learning-based shared control architecture for interactive task execution," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 329–335.

[14] M. Hagenow, E. Senft, R. Radwin, M. Gleicher, B. Mutlu, and M. Zinn, "Corrective shared autonomy for addressing task variability," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3720–3727, 2021.

[15] C. Zeng, C. Yang, H. Cheng, Y. Li, and S.-L. Dai, "Simultaneously encoding movement and semg-based stiffness for robotic skill learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1244–1252, 2020.

[16] C. Rognon, V. Ramachandran, A. R. Wu, A. J. Ijspeert, and D. Floreano, "Haptic feedback perception and learning with cable-driven guidance in exosuit teleoperation of a simulated drone," *IEEE transactions on haptics*, vol. 12, no. 3, pp. 375–385, 2019.

[17] 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)*, 2019, pp. 3630–3637.

[18] H. Latifee, A. Pervez, J.-H. Ryu, and D. Lee, "Mini-batched online incremental learning through supervisory teleoperation with kinesthetic coupling," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 5453–5459.

[19] T. Poignonec, F. Nageotte, N. Zemiti, and B. Bayle, "Simultaneous haptic guidance and learning of task parameters during robotic teleoperation," in *International Conference on Robotics and Automation (ICRA 2021)*, 2021, pp. 3619–3625.

[20] G. Li, F. Caponetto, E. Del Bianco, V. Katsageorgiou, I. Sarakoglou, and N. G. Tsagarakis, "A workspace limit approach for teleoperation based on signed distance function," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5589–5596, 2021.

[21] C. Yang, C. Zeng, P. Liang, Z. Li, R. Li, and C.-Y. Su, "Interface design of a physical human–robot interaction system for human impedance adaptive skill transfer," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 329–340, 2018.

[22] C. Zeng, X. Chen, N. Wang, and C. Yang, "Learning compliant robotic movements based on biomimetic motor adaptation," *Robotics and Autonomous Systems*, vol. 135, p. 103668, 2021.

[23] K. L. Schwaner, D. Dall'Alba, P. T. Jensen, P. Fiorini, and T. R. Savarimuthu, "Autonomous needle manipulation for robotic surgical suturing based on skills learned from demonstration," in *2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)*, 2021, pp. 235–241.

[24] C. Meeker, M. Haas-Heger, and M. Ciocarlie, "A continuous teleoperation subspace with empirical and algorithmic mapping algorithms for nonanthropomorphic hands," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 1, pp. 373–386, 2022.

[25] W. Si, N. Wang, and C. Yang, "Composite dynamic movement primitives based on neural networks for human–robot skill transfer," *Neural Computing and Applications*, pp. 1–11, 2021.

[26] C. Zeng, H. Su, Y. Li, J. Guo, and C. Yang, "An approach for robotic leaning inspired by biomimetic adaptive control," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1479–1488, 2021.

[27] 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.

[28] M. Saveriano, F. Franzel, and D. Lee, "Merging position and orientation motion primitives," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 7041–7047.

[29] Z. Lu, N. Wang, and C. Yang, "A constrained dmps framework for robot skills learning and generalization from human demonstrations," *IEEE/ASME Transactions on Mechatronics*, vol. 26, pp. 3265 – 3275, 2021.

[30] C. Yang, C. Chen, W. He, R. Cui, and Z. Li, "Robot learning system based on adaptive neural control and dynamic movement primitives," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 3, pp. 777–787, 2018.

[31] N. Hogan, "Impedance control: An approach to manipulation: Part ii—implementation," vol. 107, no. 2, pp. 8–16, 1985.

[32] C. Zeng, C. Yang, and Z. Chen, "Bio-inspired robotic impedance adaptation for human-robot collaborative tasks," *Science China Information Sciences*, vol. 63, no. 7, pp. 1–10, 2020.

[33] A. G. Marin and R. Weitschat, "Unified impedance and hybrid force-position controller with kinestatic filtering," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 3353–3359.

[34] S. Sidhik, M. Sridharan, and D. Ruiken, "Towards a framework for changing-contact robot manipulation," *arXiv preprint arXiv:2106.10969*, 2021.