

Intent-Aware Predictive Haptic Guidance and its Application to Shared Control Teleoperation

Kim Tien Ly¹, Mithun Poozhiyil¹, Harit Pandya², Gerhard Neumann³, Ayse Kucukyilmaz^{1,4}

Abstract—This paper presents a haptic shared control paradigm that modulates the level of robotic guidance, based on predictions of human motion intentions. The proposed method incorporates robot trajectories learned from human demonstrations and dynamically adjusts the level of robotic assistance based on how closely the detected intentions match these trajectories. An experimental study is conducted to demonstrate the paradigm on a teleoperated pick-and-place task using a Franka Emika Panda robot arm, controlled via a 3D Systems Touch X haptic interface. In the experiment, the human operator teleoperates a remote robot arm while observing the environment on a 2D screen. While the human teleoperates the robot arm, the objects are tracked, and the human's motion intentions (e.g., which object will be picked or which bin will be approached) are predicted using a Deep Q-Network (DQN). The predictions are made considering the current robot state and baseline robot trajectories that are learned from human demonstrations using Probabilistic Movement Primitives (ProMPs). The detected intentions are then used to condition the ProMP trajectories to modulate the movement and accommodate changing object configurations. Consequently, the system generates adaptive force guidance as weighted virtual fixtures that are rendered on the haptic device. The outcomes of the user study, conducted with 12 participants, indicate that the proposed paradigm can successfully guide users to robust grasping configurations and brings better performance by reducing the number of grasp attempts and improving trajectory smoothness and length.

I. INTRODUCTION

Collaborative robotics is an emerging field, thanks to its capability of putting robots in close distance to humans, and thus boosting the range of applications across domains. For example, significant contributions can be foreseen in enabling robots in hazardous environments [1], in minimally invasive surgery [2], operating in space [3] and for underwater exploration [4]. Although such robots have transformative capability by delivering beyond human performance through highly precise, rapid and stable movements in far more degrees of freedom than the human hand, fully autonomous solutions are not feasible due to environmental uncertainties and complex decision requirements of the tasks. This brings forward a need for developing shared control systems [5], [6], [7], where both humans and robots may benefit from one another's capabilities by dynamically and appropriately

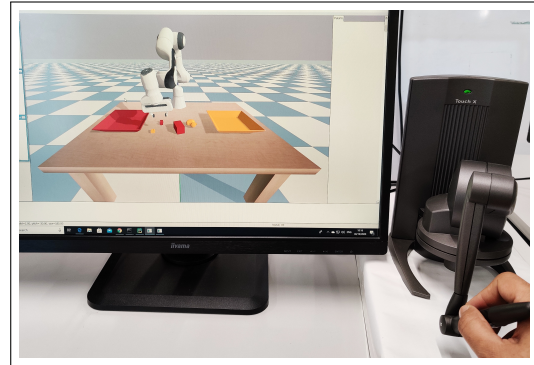


Fig. 1. The human teleoperates the Panda arm to sort objects into bins according to color, while viewing the scene on a computer screen.

assigning roles to the human and the robot [8], [9]. For example, shared control can be applied on a teleoperated nuclear waste sorting scenario, in which the human may be given the control authority to decide how the heap should be manipulated, whereas the robot can support the human by completing mechanical actions in a smooth and precise way.

Even though robots are not yet capable of matching human abilities in real world scenarios, in recent years, we witnessed a rise of robot learning methodologies, allowing robots to be programmed to learn from humans and generalise expert knowledge on motor tasks [10], [11], [12]. The current study leverages robot learning to integrate human intent awareness with robot kinematic planning within a shared-control context. We propose a haptic guidance framework to predict human intent, and based on this prediction, dynamically modulate the level of assistance to enact robot behavior for reaching dynamic objects and placing them in static bins. Human and robot autonomy are combined to appropriately guide teleoperators, so that no excessive or misleading constraints are put on human operation. This is done through an effective blend of human and robot capabilities, where the robot continuously monitors human operation to infer about the intended targets, and guides the human toward feasible robot trajectories.

In order to study the proposed paradigm, we created a teleoperated pick and place scenario. Figure 1 shows the human's view of the scene, where a haptic device is used to teleoperate a simulated robot arm, while the human observes the scene through a screen interface. Robot trajectories are pre-learned using Probabilistic Movement Primitives (ProMPs) to program a library of robot behaviours, which are then conditioned to modulate the motion to changing

¹K.T.L., M.P. and A.K. are with the School of Computer Science, University of Nottingham, UK {psxt13,Mithun.Poozhiyil,Ayse.Kucukyilmaz}@nottingham.ac.uk

²H.P. is with Toshiba Research Europe, UK harit.pandya@crl.toshiba.co.uk

³G.N. is with Autonomous Learning Robots, KIT, Germany gerhard.neumann@kit.edu

⁴A.K. is also with Lincoln Centre for Autonomous Systems (L-CAS), University of Lincoln, UK

object configurations. A Deep Q-Network (DQN) is trained to predict the human intent based on the robot’s end-effector pose. DQN is used to streamline the intent detection and generate rewards in real-time to modulate the level of robot assistance. Once the human intent is detected, the robot reactively responds to the intent, which is transferred to the human in the form of kinesthetic guidance. This intent-aware guidance paradigm allows humans the flexibility to freely move around the scene and change targets as they desire. In comparison with manual operation, the proposed method improves task performance by significantly reducing the number of grasp attempts when picking objects and improving motion smoothness and trajectory length without incurring extra energy.

II. RELATED WORK

The detection of human intention during physical interaction is not a straightforward problem. In this work, we adopt a high level definition and consider intent as a target configuration that a human aims to reach. This definition is applicable to many robotic manipulation tasks, including teleoperated picking and placing of objects. Previously, Goodrich and Jr proposed a memory-based approach to infer about user’s goals based on their intent history [13]. However, this approach depended greatly on data, which is hard to gather in interactive robotic tasks and may be affected by the personal perspectives when labelling intentions.

Humans can communicate intent through haptic cues [14], although such mechanisms are hardly applied in robotics. Moertl *et al.* has demonstrated the use of force-related information to infer about immediate human plans, and used this information to arbitrate robot autonomy [15]. Similarly, Medina *et al.* investigated the human intent to change the direction of motion [16]. Aydin *et al.* demonstrated the use of a fuzzy intent estimator and integrated it via a variable impedance controller to adjust robot autonomy [17]. An intent detection scheme which uses principle of maximum entropy over trajectories was proposed in [18]. Later, this approach was modified using Bayesian filtering in a Markov model under a probabilistic framework [19] for shared-control operation. The predictions made in this method were based on distance to closest goal or based on history of user inputs towards the closest goal. Hence, these approaches may fail when the goals are too distant or too close when inferring about intentions. In contrast, our method predicts the human intent towards the closest legible trajectories to reach goals, learned directly from human demonstrations.

Another major issue in a human-robot scenario is the programming or robot behaviours. For example, in our sorting task, we have multiple activities, such as picking and moving coloured objects and placing them in a bin with the same colour. Even though this is a simple scenario, a robot needs decision making and planning. A robot, when placed in front of such a scene would not even be capable of distinguishing between objects and bins. A human, on the other hand, could quickly develop this understanding, which can be transferred to the robot by teaching skills

to handle the scene. Our intent detection method relies on learned models of previously executed task motions. We propose the combination of intent detection and Learning from Demonstration (LfD) for motion planning as an end-to-end paradigm for robot assistance in shared control human-robot systems. LfD is an easy-to-use approach to transfer human skills into robot motions through trajectory representations learned directly from human operations [20], [10]. LfD has been used as a popular method to support collaborative robots in various industrial applications [21], [22]. Popular LfD approaches use trajectory representations by Gaussian Mixture Regression (GMR) [23], Gaussian processes (GPs) [24] or Movement Primitives (MPs). Among these, MPs have greater potential to generalize motions to different goals using concise and simple representations with small time and memory complexity. Due to the simplicity of representation and adaptation power, in our work we use MPs known as Probabilistic Movement Primitives (ProMPs) [25] to represent trajectories as weighted virtual fixtures to guide human motion, while adapting to dynamic object motion in the scene. We address the intent detection using reinforcement learning (RL) to select between multiple trajectories by computing a reward to match the nearest trajectory to the current robot end-effector pose. The use of RL enables faster matching in comparison to exhaustive search over all the phase space of all trajectories (minimising the distance to each phase of each trajectory), especially with a large number of objects. Combining artificial neural networks and RL, Deep Q-Networks (DQN) offer a model-free goal-oriented approach to solve RL problems in continuous environments.

The benefits from the mental abilities of a human operator, and the precision and reliability of a robot system can be simultaneously achieved through shared control in a teleoperation scenario [26], [27]. Recently, [28], [29] investigated shared control through optimization techniques to infer user’s goals, in turn, to provide assistance. A teleoperation shared-control architecture utilising learned trajectory properties for assistance is described in [30] for reaching tasks, however, without considering the intent of the operator. As demonstrated [6], intent recognition can be effectively used to choose the appropriate level of shared control autonomy. Similarly, the choice of control autonomy in our current work is framed as a role distribution problem [31], [32], where human and robot autonomy are balanced by choosing an appropriate level of robot autonomy, meanwhile selecting the best way to guide the human. Recently, Ewerton *et al.* introduced a framework using a combination of Gaussian mixture models (GMM) and ProMPs, to infer best trajectories when reaching for a specific target [33]. This study used RL to model virtual guidance fixtures as potential fields matching the operator’s intent to reach a single known goal. Our work is essentially different as we do not learn alternative ways to reach a target, but in contrast, we select between multiple trajectories to reach a number of targets in a complex environment, also accommodating changing object configurations. In addition, our work enables dynamic autonomy in a novel way by weighting the guidance trajectories

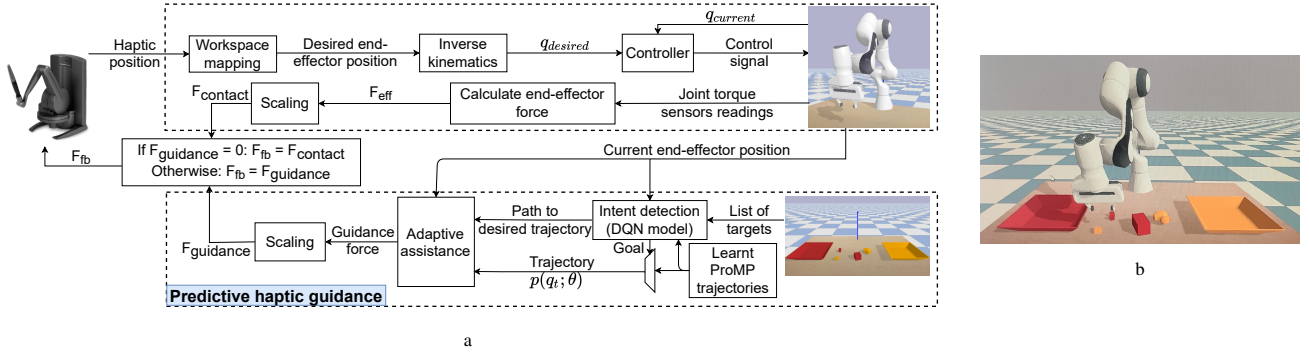


Fig. 2. a) The overview of the proposed framework to enable haptic control of the Panda arm. The bottom block presents the components of the predictive haptic guidance mechanism including trajectory planning, intent detection and adaptive guidance. (b) The experimental scenario where the remote robot arm is teleoperated to sort objects according to color.

depending on the confidence of prediction for the intent.

III. METHODOLOGY

A. Control architecture

Figure 2a represents an overview of the proposed paradigm. The system is designed for bilateral teleoperation, such that the remote arm will mirror the end-effector pose of the haptic control interface, while feedback forces are rendered on the haptic device and felt by the human. The upper block of Figure 2a describes the control loop where we use inverse kinematics to compute the desired robot joint angles based on the end-effector position of the haptic device. The contact force ($\mathbf{F}_{contact}$) is computed using the joint torque readings captured by the built-in torque sensors at each joint of the Franka Emika Panda arm. These torques are multiplied by the Jacobian to compute the end-effector torque \mathbf{F}_{eff} , which is scaled by a factor of 0.0002 to compute $\mathbf{F}_{contact}$, so that the forces can be safely rendered by the haptic device. The proposed predictive haptic guidance paradigm (Figure 2a, bottom block) generates appropriate guidance forces ($\mathbf{F}_{guidance}$) in a shared control context. It is composed of trajectory planning, intent detection and adaptive guidance components, which are detailed in the rest of this section.

Note that $\mathbf{F}_{contact}$ and $\mathbf{F}_{guidance}$ are never simultaneously rendered on the haptic device, as that could have created conflicts in the perceived forces. In order to avoid such conflicting forces, feedback force (\mathbf{F}_{fb}) is generated in the following manner:

$$\mathbf{F}_{fb} = \begin{cases} \mathbf{F}_{contact} & \text{if } \mathbf{F}_{guidance}=0 \\ \mathbf{F}_{guidance} & \text{otherwise} \end{cases} \quad (1)$$

B. Trajectory Planning

In this study, we use ProMPs to learn good robot trajectories from human demonstrations to reach and grasp a set of objects and sort them by color by placing them in static bins located in the scene (See Figure 1). The ProMP trajectories are learned prior to the experiments, and are considered to work as baseline robot plans, which are selected as appropriate to the estimated intent. Each trajectory is learned

through 10 demonstrations, and encodes the skill in terms of how each object shall be grasped and moved to each bin.

In order to capture the coordinated movement of the joints, each ProMP model is learned using the robot joint angles \mathbf{q}_t over time, $t \in \{0 \dots T\}$. Time-varying variance of the trajectories from multiple demonstrations are captured using basis functions. Basis function representations significantly reduce the number of parameters learned for each phase of the learned motion, which is one greatest advantage of ProMPs to speed up the learning procedure.

The time-varying basis matrix $\Psi_t \in \mathbb{R}^{D \times K D}$, where D and K are the number of DoFs and the basis functions, respectively, is defined as follows:

$$\Psi_t = \begin{bmatrix} \phi_t & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \phi_t \end{bmatrix} \quad (2)$$

Here, each $\phi_t \in \mathbb{R}^K$ is a basis vector that contains the normalized Gaussian basis functions

$$\phi_i(z_t) = \frac{\exp(-\frac{(z-c_i)^2}{2h})}{\sum_{j=1}^K \exp(-\frac{(z-c_j)^2}{2h})} \quad (3)$$

where h is the width of the basis and c_i is the center of the basis function i . The phase variable z is a monotonically increasing function of time, defined within the interval $z_t \in [0, 1]$, which enables temporal modulation of the trajectory. The basis functions ϕ_i are thus defined on the phase instead of time.

The basis function matrix ϕ_t is weighted by a vector \mathbf{w} to represent the trajectory for the joint as

$$\mathbf{q}_t = \phi_t \mathbf{w} \quad (4)$$

where $\mathbf{w} \in \mathbb{R}^K$. A Gaussian distribution over the weight vector \mathbf{w} contributes to the variance of the trajectory for each timestep, and can be formulated using parameters $\theta = \{\boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w\}$ as

$$p(\mathbf{w}; \theta) = \mathcal{N}(\mathbf{w} | \boldsymbol{\mu}_w, \boldsymbol{\Sigma}_w) \quad (5)$$

where $\boldsymbol{\mu}_w$ and $\boldsymbol{\Sigma}_w$ are the mean and covariance of the learned distribution, respectively. Since we have the affine

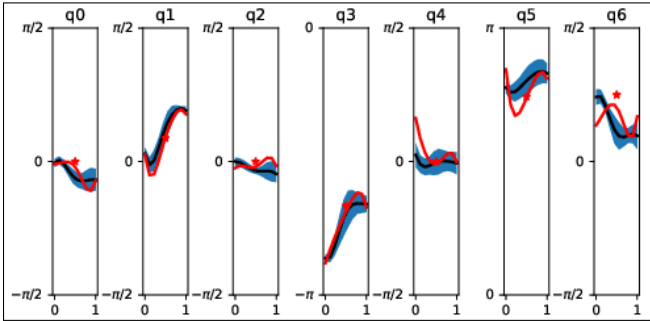
function represented in (4) for a single joint, the distribution of the states of all joints at time t using (2) can be obtained as follows (Refer to [25] for details) :

$$p(\mathbf{q}_t; \boldsymbol{\theta}) = \mathcal{N}(\mathbf{q}_t | \boldsymbol{\Psi}_t \boldsymbol{\mu}_w, \boldsymbol{\Phi}_t \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t^T) \quad (6)$$

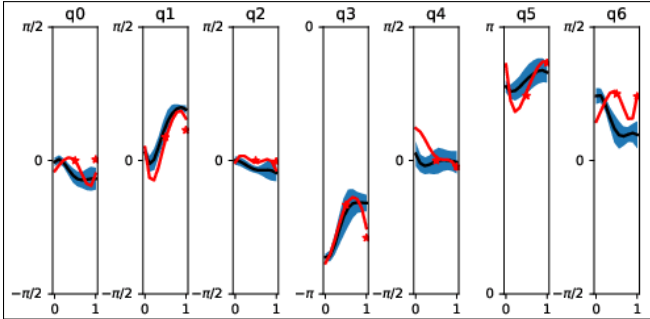
This representation enables modulation through via points using ProMP conditioning. In our context, ProMP conditioning allows the robot to modify a trajectory for passing through new via points to accommodate changes in object configurations in the scene. The modified ProMP parameters $(\boldsymbol{\mu}_w^*, \boldsymbol{\Sigma}_w^*)$ can be obtained by feeding joint angles \mathbf{q}^* to ProMP at a specific timestep t as follows:

$$\begin{aligned} \boldsymbol{\mu}_w^* &= \boldsymbol{\mu}_w + \mathbf{L}(\mathbf{q}_t^* - \boldsymbol{\Psi}_t^T \boldsymbol{\mu}_w) \\ \boldsymbol{\Sigma}_w^* &= \boldsymbol{\Sigma}_w - \mathbf{L} \boldsymbol{\Psi}_t^T \boldsymbol{\Sigma}_w \\ \mathbf{L} &= \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t (\boldsymbol{\Sigma}_q^* + \boldsymbol{\Psi}_t^T \boldsymbol{\Sigma}_w \boldsymbol{\Psi}_t)^{-1} \end{aligned} \quad (7)$$

This is a required feature as the object poses can change during manipulation, such as if pushed around while picking. The conditioning guarantees a smooth trajectory that blends with robot's current behavior and allows precise task completion as shown in Figure 3. In this example, the trajectory is conditioned to pass through two points, one at midway (Figure 3a) and another at the end (Figure 3b). The via point is the robot's joint position calculated at target end-effector positions. The plots shows that a trajectory can be smoothly altered to meet the new conditions for all joints successfully.



a) Conditioning with a new via-point, marked by an asterisk at $t = 0.5s$



b) Conditioning with a new via-point at $t = 0.5s$ and a new goal point at $t = 1s$

Fig. 3. Conditioned ProMP trajectories for seven joints (q0 - q6). The black line and the blue area refer to the mean ProMP trajectory and its variance learned through multiple demonstrations, respectively. The conditioned mean trajectory for each case is shown via the red line, whereas the red asterisks are the conditioning points.

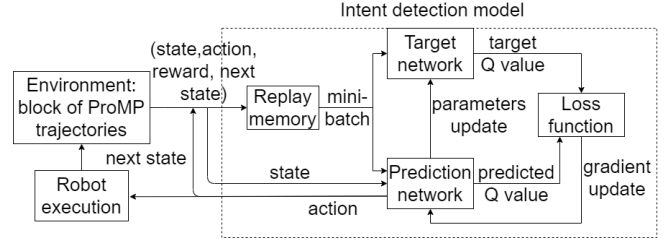


Fig. 4. Intent detection training flowchart

C. Intent Detection with DQN

This section describes how intent detection is enabled in our shared-control framework. Consider that there is a list of goals in the environment, each with a corresponding ProMP trajectory formulated in (6). To estimate the human's motion intention, we use a Deep Q-Network (DQN), which takes robot's end-effector position as input and predicts the intended ProMP trajectory. In addition, an interpolated path is computed to reach the nearest phase on the intended ProMP trajectory. This is later blended with the conditioned ProMP trajectory to compute smooth guidance. Using the DQN allows the selection of the nearest trajectory quickly, without iterating over all phases in all ProMP trajectories. Besides, the long-term reward maximization facilitates adding further constraints such as obstacles and joint limits.

The DQN uses the reward function r , which is the negative Cartesian distance between the robot's current end-effector position s and the end-effector position at the nearest phase z of the intended ProMP trajectory $p(\mathbf{q}_z; \boldsymbol{\theta})$:

$$r = - \min_{z \in [0,1]} (d(s, \mathbf{T}p(\mathbf{q}_z; \boldsymbol{\theta}))) \quad (8)$$

Here \mathbf{T} is the forward kinematics transformation, which maps the joint space to Cartesian space. Since we are dealing with multiple goals and multiple plans corresponding to each of these goals, during training, the intent recognition engine continuously loops through all possible trajectories to avoid overlapping; hence enabling convergence to a single trajectory.

The Q value is iteratively updated using a Bellman equation, which is described in (9). In this function, $r_t + \gamma \max_a Q(s_{t+1}, \mathbf{a})$ plays the role of the target value.

$$Q(s_t, \mathbf{a}_t) \leftarrow Q(s_t, \mathbf{a}_t) + \alpha [r_t + \gamma \max_a Q(s_{t+1}, \mathbf{a}) - Q(s_t, \mathbf{a}_t)] \quad (9)$$

where $Q(s, \mathbf{a})$ is the Q value for taking action \mathbf{a} at state s , α is learning rate, r is the reward value, and γ is the discount factor.

The DQN uses the loss function in (10) to perform gradient descent, where N is the number of samples, y is the Q target value and $Q(s, \mathbf{a}; \boldsymbol{\theta})$ refers to the Q value of taking action \mathbf{a} at state s given parameters $\boldsymbol{\theta}$.

$$Loss = \frac{1}{N} \sum_{i=1}^N (y_i - Q(s_i, \mathbf{a}_i; \boldsymbol{\theta}))^2 \quad (10)$$

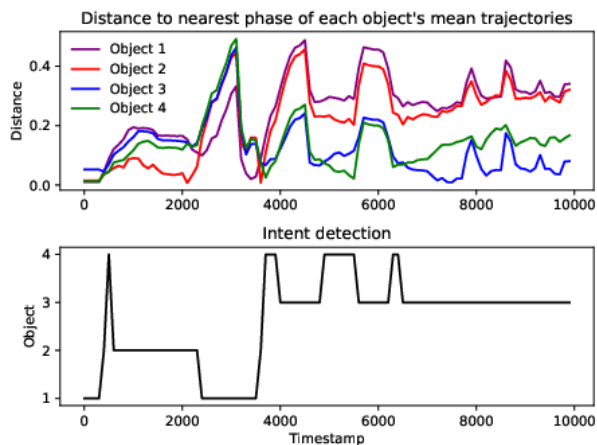


Fig. 5. Estimated intentions with respect to the distance of the end effector to nearest phase on the ProMP trajectory for reaching four targets.

Figure 4 describes the training phase of our intent detection model. The initial end-effector positions are randomly distributed in the workspace. Epsilon-Greedy algorithm balances between exploration and exploitation in choosing steps towards six directions. The episode terminates when the robot reaches one phase of any learned ProMP trajectory. We use experience replay to randomly pick samples from this memory to get the collected experience for training the network. The system contains two networks, which are similar in structure but different in parameters, known as target and prediction (also called evaluation) networks. The parameters of the prediction nets will transfer to the target one after a defined number of iterations.

Figure 5 demonstrates how estimated intent changes with respect to Cartesian distance calculation over a period of time. The results are presented on a simple scenario, where the human moves freely in between four target objects (numbered 1 – 4) within a 3D scene. Respectively, four ProMP trajectories are learned to reach each object. The DQN model can detect the target that matches the nearest trajectory. Note that due to the ProMP representation, the learned trajectory distributions may overlap and hence the resulting intent may oscillate in case two trajectories are close to each other.

The intent recognition engine continuously computes a reward during motion, based on the current end-effector position. This reward is used to select the target object and the corresponding trajectory. A sequence of steps are computed to attract the human toward the nearest target trajectory phase. Figure 6 illustrates the operation of the intent detection engine in the example scene used in Figure 5. Here, the human starts the motion from point A, and moves to reach point B along the blue trajectory. At this point, the DQN model outputs the current intention as Object 2, and estimates the nearest phase on the corresponding ProMP trajectory to reach this object. The guidance action is computed to attract the human toward this trajectory (marked with the red dotted line between B and the ProMP trajectory). However, the human doesn't comply with guidance, and

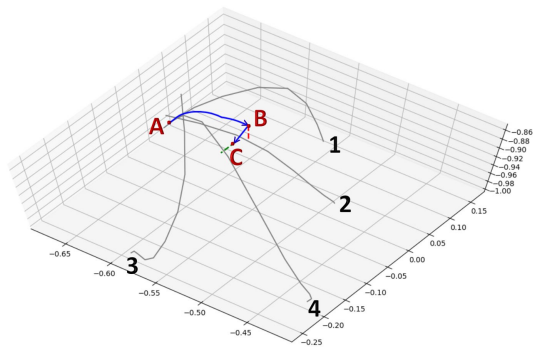


Fig. 6. Intent detection responds to the human operator's behaviors: The grey lines are the mean ProMP trajectories towards 4 different objects (1 – 4). Blue lines connecting points A, B, C indicate the movement of the human. The red and green dotted lines are the paths output by the DQN model at point B and C, respectively. The guidance attracts the human toward the most likely ProMP trajectory, however it can be easily overridden.

continues the motion to proceed to Point C along the blue trajectory. At C, the intent recognition engine estimates that the human's intent has changed to Object 4, hence guidance is provided to lead the human to the closest phase in the corresponding trajectory over the green dotted line.

In order to ensure teleoperators retain control over robotic autonomy, we implemented a mechanism to break robot guidance if the human doesn't choose to move toward the guided trajectory *e.g.* move to another direction within a given time window, (0.3s in our experiment). In this case, guidance forces are cancelled and the DQN will detect another target.

D. Adaptive Haptic Guidance

As the intended target is estimated, we generate two consecutive motion plans. The first is the output of the intent detection engine, and describes a path to snap to the closest phase on the intended ProMP trajectory. The second one is the ProMP trajectory conditioned with the (possibly changed) object pose. The robotic autonomy is programmed as haptic virtual fixtures guiding on these two trajectories. In order to avoid jerky changes in the guidance forces when starting and stopping the guidance, an exponential function is applied on the force profile to damp the forces.

As mentioned earlier, modulating the role switching behavior between robot autonomy and human control is important in shared control. If the robot is too persistent in guiding the human, the operation could become restraining for human decision taking. What is important to note here is that the intent detection engine estimates a path to the intended trajectory and this information can be used to capture the confidence about the correctness of the estimated intent. This is implemented regarding the number of steps required to snap to the estimated ProMP trajectory. Since this path leads the robot from its current position to the nearest phase on trajectory, a large number of steps indicate that the human is relatively far away from the robot's plan, thus the belief that the detected target is correct is weak. Accordingly, the adaptive guidance mechanism we propose weights the force

guidance given through the haptic device by introducing a belief coefficient based on the distance to the intended ProMP trajectory. The belief coefficient will make sure that the engine doesn't put too much force towards one trajectory while the user is still moving undecidedly. As a result, if the user suddenly switches to a different direction, the force guidance will be switched off until the next detection.

Equation 11 denotes how we calculate our adaptive force, where k is the scaling gain ($k = 9$), N is the predefined threshold for the maximum number of steps to the ProMP trajectory to consider as user intention, r is the number of steps taken to the intended phase on the ProMP trajectory, s_{t+1} is the next expected end-effector position from s_t on the computed trajectory.

$$F_{\text{guidance}} = k(N - r)(s_{t+1} - s_t) \quad (11)$$

IV. EXPERIMENTS

In order to evaluate the proposed framework, we conducted a user study in a teleoperated pick-and-place scenario. A 3D Systems Touch X haptic device is used to control a Franka Emika Panda robot arm with 2-finger grippers as the remote system. The participants monitored the operation of the remote arm through a 2D screen.

12 subjects (1 female, 11 male) aged between 23 and 38 from different academic backgrounds, participated in the study. 5 participants had no prior experience working with robotic arms and 3 had experience using haptic devices. At the beginning of the trials, the participants were given an instruction sheet providing the details of the experiment and the required task. A practice session is presented to allow participants to get familiar with the robot and teleoperation using the haptic device. The practice session was maintained until participants felt comfortable with the control interface. The participants signed an informed consent form at the beginning of the experiment. A full ethics approval (reference 2019-Jul-0802) was obtained from the Human Ethics Committee in the University of Lincoln, where the experiments took place in, and a full risk assessment was completed before the commencement of the studies. Appropriate automatic and manual safety measures were installed, including physical and software-based kill switches, and moderated by the experimenter to stop the haptic device in case of emergency.

Among 12 subjects, 2 were left-handed and all of them used their dominant hands to control the remote robot. The Franka fingers were controlled using the button on the haptic device stylus. The experimental scene is shown in Figure 2b, where four objects and two trays, colored either red or yellow, were placed within the workspace of the Panda robot. Participants were asked to sort each object to the trays with matching colors. The colors of the trays were constant, whereas the colors of objects were pseudo-randomly chosen (i.e. the same colors were selected in each trial, so all participants experimented with the same object color sequence). The controller was agnostic to the colors; as a result, only the human could decide which object should go to which tray.

The objects were located relatively close to each other, so that the intent detection was not straightforward. We tested two conditions within the experiment:

- **Manual mode:** Users had complete control of the remote arm through the haptic device. They were provided with collision forces with objects and the scene.
- **Guided mode:** In addition to collision forces, the users were also provided with guidance forces that led the haptic stylus over a learned trajectory to reach and place objects, using the proposed intent-aware predictive haptic guidance paradigm.

A balanced within-subjects design was used in the experiment, so that all participants experimented with both conditions. The subjects were randomly allocated into two groups to eliminate ordering effects, where the first group experimented with the manual mode first, whereas the second group experimented with the guided mode first. Participants were not informed in advance which mode they experimented on. Each participant completed three trials in each condition, carrying out 12 picking and placing operations (3 trials \times 4 objects in the scene) in total.

As the robot does not know which tray the objects should be sorted in, the high-level decisions were made completely by the human operators. In our experimental scenario, the predictive haptic guidance was provided during both picking and placing phases of the task. Object positions were tracked continuously to condition the ProMP trajectories. Objects that are already sorted were removed from the list of targets that were considered for intent detection. We recorded time-stamped joint torques, positions and velocities during the experiment. The following metrics are used to evaluate the results of the user study:

- **Time:** The total time to pick and place four objects is recorded for each trial.
- **Number of correctly sorted objects:** The number of correctly sorted objects, *i.e.* those placed in a tray of the same color, are counted for each trial.
- **Number of grasping attempts:** The number of times that the user pressed the haptic device button to close the Franka grippers is counted as the number of grasping attempts per trial.
- **Energy:** We compute the energy consumption of the task by integrating the human exerted power over time as $E = \sum_{t=0}^T \tau(t) \dot{q}(t)$, where τ and \dot{q} are joint torques and velocities of the remote robot, respectively, and T is the duration of the trial.
- **Trajectory length:** The total length of the trajectory in each trial is computed to measure trajectory complexity.
- **Trajectory smoothness:** Spectral arc length [34] of robot joint motions is computed, and their average is used to measure motion smoothness at each trial.

At the end of each condition block, the participants were given the NASA Task Load Index (TLX) [35] to assess their perceived workload on a 5-point Likert scale.

In order to investigate the statistical significance of observed differences between the conditions, we conducted t-

tests for the continuous level measures. Wilcoxon signed-rank tests were conducted for ordinal level survey data.

Figures 7 and 8 illustrate a comparison of the quantitative metrics. The results indicate no significant effect of the conditions (manual vs. guided mode) on task completion time and energy. On the other hand, a statistically significant difference is observed in terms of the trajectory smoothness ($t(36) = 3.01, p < 0.005$) and trajectory length ($t(36) = 2.27, p < 0.05$). This indicates that the subjects performed significantly more smoothly under guidance, and completed the motions in a more targeted and concise manner.

In terms of task performance, we observed that the subjects completed the sorting task almost perfectly (with the exception of a single object being misplaced in the manual mode over all trials) in both conditions. A statistical significant difference between conditions is observed in the number of grasping attempts with $t(36) = 2.02$ and $p < 0.05$.

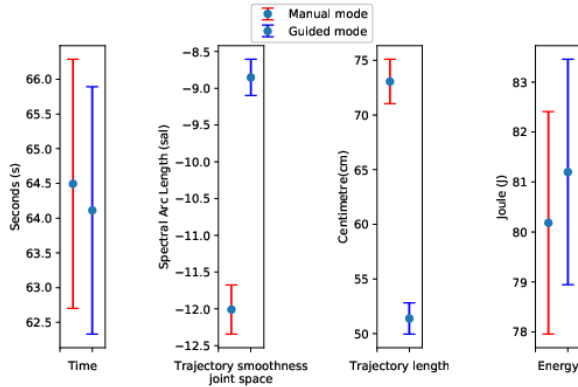


Fig. 7. Mean values and standard errors of the means for task completion time, trajectory smoothness in joint space, trajectory length, and energy, under manual and guided modes. For the spectral arc length metric, the values closer to zero indicate smoother trajectories.

Figure 9 shows the results of the workload assessments under each condition. No statistical significant differences are observed for any of the subjective measures. Overall, the participants assessed the workload rather low in all dimensions of the NASA-TLX in both modes, indicating that the task was not perceived as hard or mentally demanding.

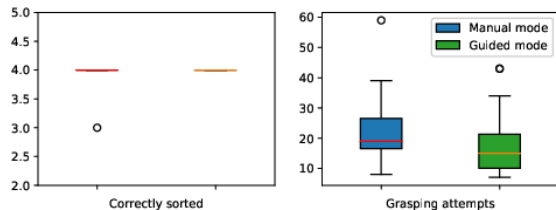


Fig. 8. The medians and IQRs for the number of correctly sorted objects and grasping attempts under each condition.

V. CONCLUSIONS AND FUTURE WORK

This study proposes a predictive haptic guidance methodology, combining intent detection, trajectory planning and

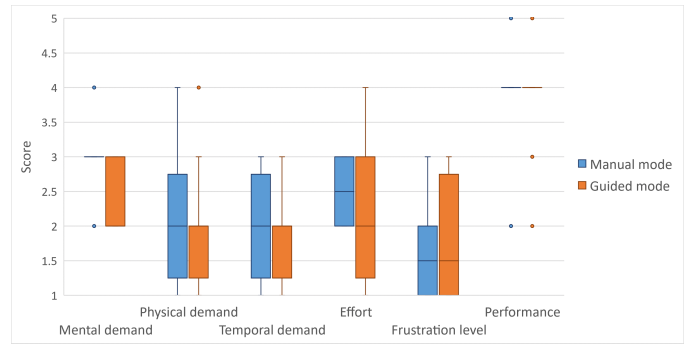


Fig. 9. Comparison of perceived task workload in two modes.

adaptive assistance as a shared-control solution. This is an end-to-end mechanism, applicable to teleoperated pick and place tasks in industrial applications. We employed a deep reinforcement learning method for detecting an operator's intended goal and finding the shortest path towards baseline ProMP trajectories, learned from experts' demonstrations. Attraction toward dynamic objects are handled leveraging the Gaussian conditioning property of ProMPs. Force guidance is adaptively rendered on the haptic device, based on the robot's confidence of the detected intent.

The results of the user study shows that our system can provide operators with intuitive assistance. With force guidance, participants can grasp objects more precisely and optimize their trajectories, which is manifested in significantly lower number of grasping attempts, shorter trajectory length and smoother joint trajectories.

The proposed framework has the ability to detect users' intention in real time and generate a complete trajectory from current pose to the intended target based on the learned knowledge from experts' demonstration. The current study works with a limited number of trajectories, hence its ability to work in unknown scenes is not readily demonstrated. Please note that the paradigm can be extended to iteratively learn trajectories from experience to handle situations where current robot trajectories are far away from learned models. However, this is beyond the scope of the current study and will be examined as part of future work.

Due to COVID-19 and related workplace restrictions, the experimental study was conducted on a simulation framework, where manipulating and grasping with robot is easier than what one would have with a real robot arm. Through using the haptic interface, the participants were able to feel realistic sensations, and strong significant differences are observed during the experiment. As future work, we will implement the interaction mechanism on a physical robot.

The current paradigm does not consider strong conflicts between the plans of the user and the robot. Although the framework can handle slight changes to target configurations, larger diversions are not tested in this experiment. In addition, the paradigm does not model situations where human intentions are significantly different than the plans in the robot's repertoire. To tackle this, active learning mechanisms can be integrated in the current framework to enable iterative

learning to enrich the robot’s behaviour library. In addition, extending on our recent work [36], [37], we plan to integrate active conflict recognition within this framework, and as a result, handle not only collaborative but also conflicting scenarios in physical human-robot interaction.

ACKNOWLEDGMENT

This work is supported by CHIST-ERA and EPSRC (“HEAP”, EP/S033718/2; “NCNR” Flexible Partnership Funding “CoRSA”, EP/R02572X/1: 742782).

REFERENCES

- [1] K. Qian, A. Song, J. Bao, and H. Zhang, “Small teleoperated robot for nuclear radiation and chemical leak detection,” *International Journal of Advanced Robotic Systems*, vol. 9, no. 3, p. 70, 2012. [Online]. Available: <https://doi.org/10.5772/50720>
- [2] A. J. Madhani, G. Niemeyer, and J. K. Salisbury, “The black falcon: a teleoperated surgical instrument for minimally invasive surgery,” in *Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovations in Theory, Practice and Applications (Cat. No.98CH36190)*, vol. 2, 1998, pp. 936–944 vol.2.
- [3] T. Imaida, Y. Yokokohji, T. Doi, M. Oda, and T. Yoshikawa, “Ground-space bilateral teleoperation of ets-vii robot arm by direct bilateral coupling under 7-s time delay condition,” *IEEE Transactions on Robotics and Automation*, vol. 20, no. 3, pp. 499–511, 2004.
- [4] Qingping Lin and Chengli Kuo, “Virtual tele-operation of underwater robots,” in *Proceedings of International Conference on Robotics and Automation*, vol. 2, 1997, pp. 1022–1027 vol.2.
- [5] D. Abbink, M. Mulder, and E. Boer, “Haptic shared control: smoothly shifting control authority?” *Cognition, Technology and Work*, vol. 14, 03 2012.
- [6] A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, “Conveying intentions through haptics in human-computer collaboration,” in *2011 IEEE World Haptics Conference*, 2011, pp. 421–426.
- [7] A. Kucukyilmaz, S. O. Oguz, T. M. Sezgin, and C. Basdogan, “Improving human-computer cooperation through haptic role exchange and negotiation,” in *Immersive Multimodal Interactive Presence*. Springer, 2012, pp. 229–254.
- [8] N. Jarrasse, V. Sanguineti, and E. Burdet, “Slaves no longer: review on role assignment for human-robot joint motor action,” *Adaptive Behavior*, vol. 22, no. 1, pp. 70–82, 2014.
- [9] D. Gopinath, S. Jain, and B. D. Argall, “Human-in-the-loop optimization of shared autonomy in assistive robotics,” *IEEE Robotics and Automation Letters*, vol. 2, no. 1, pp. 247–254, 2016.
- [10] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, “Survey: Robot programming by demonstration,” *Handbook of robotics*, vol. 59, no. BOOK.CHAP, 2008.
- [11] P. Pastor, H. Hoffmann, T. Asfour, and S. Schaal, “Learning and generalization of motor skills by learning from demonstration,” in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 763–768.
- [12] M. P. Deisenroth, G. Neumann, J. Peters *et al.*, “A survey on policy search for robotics,” *Foundations and trends in Robotics*, vol. 2, no. 1-2, pp. 388–403, 2013.
- [13] M. Goodrich and D. Jr, “Seven principles of efficient human robot interaction,” vol. 4, 11 2003, pp. 3942 – 3948 vol.4.
- [14] N. Stefanov, A. Peer, and M. Buss, “Role determination in human-human interaction,” in *World Haptics 2009-Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, 2009, pp. 51–56.
- [15] A. Mörtl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, “The role of roles: Physical cooperation between humans and robots,” *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1656–1674, 2012.
- [16] J. R. Medina, M. Lawitzky, A. Molin, and S. Hirche, “Dynamic strategy selection for physical robotic assistance in partially known tasks,” in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 1180–1186.
- [17] Y. Aydin, N. Arghavani, and C. Basdogan, “A new control architecture for physical human-robot interaction based on haptic communication,” in *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2014, pp. 122–123.
- [18] A. D. Dragan and S. S. Srinivasa, “A policy-blending formalism for shared control,” *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013.
- [19] S. Jain and B. Argall, “Probabilistic human intent recognition for shared autonomy in assistive robotics,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, no. 1, pp. 1–23, 2019.
- [20] S. Schaal, “Is imitation learning the route to humanoid robots?” *Trends in Cognitive Sciences*, vol. 3, pp. 233–242, 06 1999.
- [21] S. El Zaatari, Y. Wang, Y. Hu, and W. Li, “An improved approach of task-parameterized learning from demonstrations for cobots in dynamic manufacturing,” *Journal of Intelligent Manufacturing*, pp. 1–17, 2021.
- [22] L. Roveda, M. Magni, M. Cantoni, D. Piga, and G. Bucca, “Human-robot collaboration in sensorless assembly task learning enhanced by uncertainties adaptation via bayesian optimization,” *Robotics and Autonomous Systems*, vol. 136, p. 103711, 2021.
- [23] S. Calinon, F. D’halluin, E. L. Sauser, D. G. Caldwell, and A. G. Billard, “Learning and reproduction of gestures by imitation,” *IEEE Robotics Automation Magazine*, vol. 17, no. 2, pp. 44–54, 2010.
- [24] M. P. Deisenroth, D. Fox, and C. E. Rasmussen, “Gaussian processes for data-efficient learning in robotics and control,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 2, pp. 408–423, 2013.
- [25] A. Paraschos, C. Daniel, J. Peters, and G. Neumann, “Using probabilistic movement primitives in robotics,” *Autonomous Robots*, vol. 42, 07 2017.
- [26] D. Saraphis, V. Izadi, and A. Ghasemi, “Towards explainable cobots: Developing confidence-based shared control paradigms,” in *Dynamic Systems and Control Conference*, vol. 84270. American Society of Mechanical Engineers, 2020, p. V001T17A006.
- [27] M. J. Zeestraten, I. Havoutis, and S. Calinon, “Programming by demonstration for shared control with an application in teleoperation,” *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1848–1855, 2018.
- [28] S. Javdani, H. Admoni, S. Pellegrinelli, S. S. Srinivasa, and J. A. Bagnell, “Shared autonomy via hindsight optimization for teleoperation and teaming,” *The International Journal of Robotics Research*, vol. 37, no. 7, pp. 717–742, 2018.
- [29] M. Laghi, M. Maimeri, M. Marchand, C. Leparoux, M. Catalano, A. Ajoudani, and A. Bicchi, “Shared-autonomy control for intuitive bimodal tele-manipulation,” in *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2018, pp. 1–9.
- [30] 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)*. IEEE, 2017, pp. 329–335.
- [31] S. O. Oguz, A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, “Haptic negotiation and role exchange for collaboration in virtual environments,” in *2010 IEEE Haptics Symposium*. IEEE, 2010, pp. 371–378.
- [32] P. Evrard and A. Kheddar, “Homotopy switching model for dyad haptic interaction in physical collaborative tasks,” in *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2009, pp. 45–50.
- [33] M. Ewerton, O. Arenz, and J. Peters, “Assisted teleoperation in changing environments with a mixture of virtual guides,” *Advanced Robotics*, pp. 1–14, 07 2020.
- [34] S. Balasubramanian, A. Melendez-Calderon, A. Roby-Brami, and E. Burdet, “On the analysis of movement smoothness,” *Journal of neuroengineering and rehabilitation*, vol. 12, no. 1, p. 112, 2015.
- [35] S. G. Hart, “Nasa-task load index (NASA-TLX); 20 years later,” in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9. Sage publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.
- [36] A. Kucukyilmaz and I. Issak, “Online identification of interaction behaviors from haptic data during collaborative object transfer,” *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 96–102, 2019.
- [37] Z. Al-Saadi, D. Sirintuna, A. Kucukyilmaz, and C. Basdogan, “A novel haptic feature set for the classification of interactive motor behaviors in collaborative object transfer,” *IEEE Transactions on Haptics*, vol. 14, no. 2, pp. 384–395, 2021.