Waypoints Updating based on Adam and ILC for Path Learning in Physical Human-Robot Interaction

Jingkang Xia, Chenjian Song, Deqing Huang, Xueyan Xing, Lei Ma, Yanan Li

***Abstract*—This paper presents a novel method for learning and tracking of the desired path of the human partner in physical human-robot interaction. Combining the Adam optimization algorithm with iteration learning control (ILC), a path learning method is designed to generate and update reference waypoints according to the human partner's desired path. This method firstly uses the Adam optimization algorithm to update the robot's reference waypoints in an online manner. Then, an ILC is developed to further modify the waypoints and reduce the difference between the robot's actual path and the human partner's desired path in an iterative manner. Simulations and experiments on a 7-DOF Sawyer robot are carried out to show the effectiveness of our proposed method.**

***Index Terms*—Physical human-robot interaction; Iterative learning control; Waypoints optimization.**

I. Introduction

Human-robot interaction (HRI) combines human flexibility with the repeatability and high precision of robots, thereby effectively reducing human workload and improving work efficiency [1], [2]. It can be used in a variety of application domains such as collaborative assembly, heavy load transport and robot-assisted rehabilitation. A typical HRC scenario is shown in Fig.1, where a robotic arm needs to move along the path from P_0 to P_N, guided by a human hand through physical interaction. As the human partner's desired path is unknown to the robot and is subject to uncertainties, how to design the robot controller to make it efficiently follow the human partner is still an open problem.

In early works of physical human-robot interaction (pHRI), an impedance controller is designed to make the robot passively follow the human partner, and then the recorded movement is played back, which is known as programming by demonstration (PbD) [3]–[5]. In [4], a method for human users to teach the robotic arm a movement path based on the interaction force is proposed, which can effectively ensure the safety of human users during the teaching and playback process. [5] proposes a method based on teaching force shaping to solve the problem of insufficient accuracy when the contact force is small in the teaching process. PbD effectively simplifies the operation of robotic arms and promotes their use in applications typical in small and medium-sized enterprises. However, the traditional PbD method requires multiple demonstrations to find a suitable task path for the robot. Moreover, the nature of offline demonstrations limits the generalizability of the PbD method, which is crucial when the robot operates in an uncertain environment.

The work was partially supported by the National Natural Science Foundation of China under Grants 61773323, 61433011, 61603316, 61733015, U1934221 and UK EPSRC grant EP/T006951/1. (Jingkang Xia and Chenjian Song contributed equally to the work.) (Corresponding author: Yanan Li.)

J. Xia, C. Song, D. Huang and L. Ma are with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, P. R. China.

X. Xing and Y. Li are with the Department of Engineering and Design, University of Sussex, Brighton BN1 9RH, UK. email: yl557@sussex.ac.uk

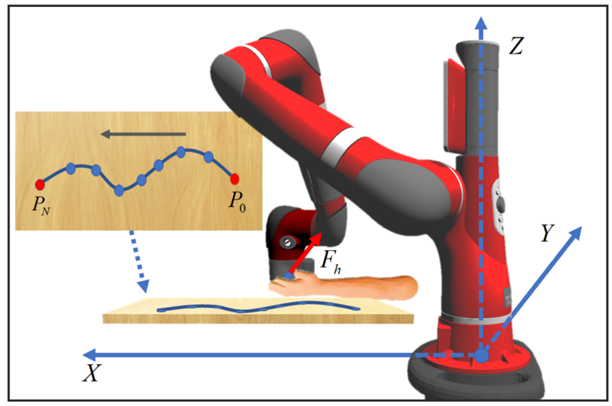

Fig. 1: A typical human-robot interaction (HRI) scenario, where a human hand guides the endpoint of a robotic manipulator to follow the contour of a workpiece. This contour is unknown to the robotic manipulator, represented by a number of waypoints from P_0 to P_N. F_h is the interaction force between the human partner's hand and the robotic manipulator.

Considering the repetitiveness of many HRI applications, iterative learning control (ILC) can be used in PbD tasks, which was initially developed for motion control of systems performing repetitive operations such as in [6]–[10]. The idea of the ILC method is to improve a system's performance by learning its periodic characteristics (see [11] for a review). However, due to various uncertainties, the assumption of periodicity required by ILC may be invalid, so there are research works dealing with varying periods. In [12], an ILC scheme with an iteration-average operator is proposed for discrete-time linear systems. It is further developed in [13] for nonlinear dynamic systems with random changes in the iteration period. In our previous work [14], we propose a method based on spatial iterative learning to learn the desired path of the robot's human partner, which is assumed to be periodic in space instead of in time. In [15], we develop a period-varying ILC scheme for HRI where the human movement is assumed to be periodic but with uncertain time durations. Despite these research efforts, the existing ILC schemes are based on a strong assumption of either a spatial or temporal period, which may be violated due to highly uncertain human movements.

Based on the Adam optimization method in [16], a waypoint update algorithm is developed in [17] to improve the physical interaction between a robot and a human user during the process of robot-assisted dressing. An interesting point of the algorithm in [17] is that the Adam optimization method generates and updates the robot's reference waypoints instead of a predefined trajectory or a path. Following this idea, we will develop a novel ILC to update the robot's reference waypoints according to the interaction force during the movement in the previous iteration, so as to realize the learning of the human's desired path. This ILC will be combined with the Adam method, which will be implemented in each individual iteration, leading to two-level path learning: one is iterative and the other is continuous in time. The proposed algorithm will be compared with the existing ILC method and the Adam method, showing its advantages in terms of faster learning and reduced human effort.

Based on the above discussions, we highlight the following contributions of this paper:

1) A path optimization and learning algorithm is developed for updating the robot's waypoints based on the interaction force, enabling the robot to adapt to different human movements during HRI.
2) The proposed method does not require the assumption of repetitive task with a certain period, and addresses the limitation of the ILC method in [14].
3) By combining the improved Adam path optimization algorithm with ILC, our method achieves faster learning and reduced human effort, compared to the algorithm in [17].

The rest of this paper is organized in the following order. Problem formulation is given in Section II. The details of the proposed method and analysis of the system performance are explained in Section III. Simulation and experimental results are presented in Sections IV and V, respectively. Finally, conclusions are drawn and possible future works are suggested in Section VI.

II. Problem Formulation

In this paper, we consider a typical human-robot interaction scenario that is composed of a robotic manipulator and a human arm. The human arm guides the robotic manipulator to complete a path following task, e.g. following the human partner's desired path as shown in Fig. 1. The path is determined by the contour or surface of the workpiece and unknown to the robotic manipulator. The interaction force between the human hand and the robotic manipulator is measured by a force sensor at the end-effector of the robotic manipulator.

A. System's dynamic model

It is often desirable to describe the robot's dynamics in the Cartesian space for the convenience of analysis, when the interaction takes place at the end-effector. The robot dynamics in the Cartesian space are given by

$$H_x\ddot{X} + C_x\dot{X} + G_x = J^{-T}u + F_h \tag{1}$$

where $X, \dot{X}, \ddot{X} \in \mathbb{R}^n$ represent the robot's position, velocity and acceleration vectors of the end-effector, respectively; $H_x \in \mathbb{R}^{n\times n}$ is the symmetric positive definite mass matrix; $C_x\dot{X} \in \mathbb{R}^n$, $G_x \in \mathbb{R}^n$ denote the centrifugal force and gravity, respectively; $u \in \mathbb{R}^n$ is the joint torque applied by the robot's actuators; $J \in \mathbb{R}^{n\times n}$ is the Jacobian matrix that relates the joint velocity to the linear and angular velocities of the end-effector and $F_h \in \mathbb{R}^n$ is the interaction force that can be measured by a force/torque sensor.

Let $e = X - X_r$, where X_r is the reference trajectory and e is the tracking error vector. Thus, $X = X_r + e$, $\dot{X} = \dot{X}_r + \dot{e}$, $\ddot{X} = \ddot{X}_r + \ddot{e}$. Combining these equations with Eq. (1), the tracking error dynamics can be described as

$$H_e\ddot{e} + C_e\dot{e} + G_e = F + F_h \tag{2}$$

where $H_e = H_x$, $C_e = C_x$, $G_e = H_x\ddot{X}_r + C_x\dot{X}_r + G_x$ and $F = J^{-T}u$. By designing robot controller as $F = -K_v\dot{e} - K_pe + C_e\dot{e} + G_e - F_h$, we obtain

$$H_e\ddot{e} + K_v\dot{e} + K_pe = 0 \tag{3}$$

where K_v, K_p are two positive definite matrices. Eq. (3) indicates tracking of the reference trajectory, i.e. $e = 0$ and $X = X_r$.

B. Path description

As discussed above, this paper studies a path following task in the robot's operational space, where the human partner's desired path is determined by the workpiece's contour but unknown to the robot. Therefore, the robot's control objective is different from the traditional trajectory tracking control where the reference trajectory is given. To formulate the problem under study, we introduce the following analysis.

According to human motor control [18], [19], the interaction force can be expanded as

$$F_h = K_h(X_h - X) \tag{4}$$

where K_h is the equivalent stiffness of the human arm and X_h forms the desired path of the human. It is noted that the human model (4) is used for analysis purpose only. In practice, the interaction force F_h is applied by the human user to the robot and measured by a force sensor but not generated by this model. Since the interaction force F_h can be measured, it can be used to estimate X_h which is unknown to the robot. In an ideal case, if we can achieve $F_h = 0$, it indicates that the robot's actual position $X = X_h$.

Different from the previous work [14], we will not update the robot's reference trajectory in the time domain but will introduce a waypoints updating method to generate the robot's reference path. As shown in Fig. 2, the reference path of the robot's end-effector in the t-th iteration is defined by a set of waypoints

$$W_t = \left\{P_{(1,t)}, \cdots, P_{(i,t)}, \cdots, P_{(N,t)}\right\} \tag{5}$$

where $P_{(i,t)}$ is one of the waypoints in the t-th iteration and $P_{(i,t)} = [x_i, y_i, z_i]_t$, N is the number of the waypoints in the t-th iteration. The desired path of the human partner is defined by W_{tar} and $P_{(i,tar)}$ is one of the desired waypoints in W_{tar}. The objective of the proposed method is to reduce the error between $P_{(i,tar)}$ and $P_{(i,t)}$, and make W_t close to W_{tar}. In the following section, we will introduce the proposed waypoints updating method in details.

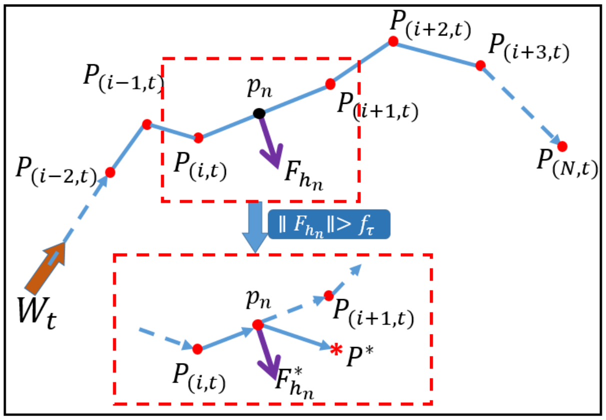

Fig. 2: W_t is a set of waypoints that defines the robot's reference path in the t-th iteration, where $P_{(i,t)}$ is one of the waypoints, p_n is a path point and F_{h_n} is the detected interaction force at this point. When $\| F_{h_n} \| > f_\tau$ where f_τ is a waypoints updating threshold, a new waypoint P^* is obtained and $F^*_{h_n}$ is set as F_{h_n}.

III. Waypoints Update

This section dedicates to designing the waypoints updating scheme that includes the Adam optimization and the ILC. As is shown in Fig. 3, the whole control framework consists of three units: robot control, Adam optimization and ILC update. While robot control has been briefly introduced in the previous section, we will present the details of the Adam optimization and ILC update methods in the following two subsections, respectively.

A. Adam optimization

In order to plan the robot's motion to follow the movement of the human partner during the interaction, the Adam optimization algorithm in [17] is used to update the robot's reference waypoints in real time according to the interaction force between human and robot.

As introduced above, W_t is a set of reference waypoints in the t-th iteration. For any adjacent waypoints $P_{(i,t)}$ and $P_{(i+1,t)}$, we set them as the start point P_{start} and end point P_{end}, respectively. The motion planning for P_{start} and P_{end} is executed using the path planning library (of the Sawyer robot in this paper) such that the path points $p_n, n \in \mathbf{N}^+$ between P_{start} and P_{end} are obtained. The interaction force F_{h_n} at each path point is measured by the force sensor, as is shown in Fig. 2.

If $\| F_{h_n} \| > f_\tau$ where f_τ is a waypoints updating threshold, a new waypoint P^* is added to the existing set of waypoints (how to compute P^* will be introduced later). Then we set p_n and P^* as the new P_{start} and P_{end}, and execute the path planning to obtain new path points between them. The interaction force is measured continuously with the new path, until $|\overrightarrow{p_n P^*}| \leq R$, i.e. the distance from p_n to P^* is less than R which is set to be small. At last, we set the final P^* as $P^*_{(i+1,t)}$ and let the robot move to the next waypoint $P_{(i+2,t)}$.

In the following, we explain how to compute the new waypoint P^*. For this purpose, a force function in the t-th iteration is designed as

$$\varepsilon_t + = \begin{cases} \| F^*_{h_n} \|, & \| F_{h_n} \| \geq f_\tau \\ 0, & \| F_{h_n} \| < f_\tau \end{cases} \tag{6}$$

As the objective is to make the interaction force $F_h = 0$, the idea of the Adam optimization is to make $\varepsilon_t = 0$ in the t-th iteration. According to the biased moment estimate of the Adam optimization method, the biased first and second moment estimates of F_h are defined as

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\mathbf{E} F_h \tag{7}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\mathbf{E} F_h^2 \tag{8}$$

where $\beta_1 \in [0 \quad 0.9]$ and $\beta_2 \in [0.99 \quad 0.999]$ are hyper-parameters which control the exponential decay rates of the moment estimates and $\mathbf{E}$ is the unit matrix. β_1 and β_2 are set with a low value when reliable new data is available; otherwise, they are set with a high value (more detailed analysis can be found in [16]). Since m_t and v_t are initialized to zero matrix, the moment estimates will be biased towards zero, especially during the first few iterations. Therefore, here we take the same strategy as in the Adam optimization method to use the bias-corrected estimates $\hat{m}_t$ and $\hat{v}_t$, which are given below

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{9}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{10}$$

By defining the updated waypoint $P^* = [P^*_x, P^*_y, P^*_z]$ that represents how the robot locally adjusts its position based on the force information and referring to the Adam optimization algorithm [17], we can set the $x-$component P^*_x as below

$$P^*_x = P_{end_x} + \frac{\alpha \hat{m}_{t_x}}{\sqrt{\hat{v}_{t_x}} + \epsilon} \tag{11}$$

where $\alpha > 0$ is the updating rate and $\epsilon > 0$ is a denominator correction term to avoid denominator being too close to 0. P^*_y and P^*_z can be obtained in a similar way as computing P^*_x.

B. ILC update

In this subsection, the path point W^*_t generated by using the Adam optimization is further updated through the ILC method. As the repetitiveness of the human movement is not guaranteed in different iterations, we develop a novel ILC algorithm based on spatial waypoints.

From the human motor control model in Eq. (4), it can be found that the interaction force can be used to represent the position difference between the actual path point of the robot and the desired path point of the human partner. At the end of each iteration, we can obtain a set of waypoints $W^*_t = \left\{ P_{(1,t)}, \cdots, P^*_{(i,t)}, P^*_{(i+1,t)}, P^*_{(i+2,t)}, \cdots, P_{(N,t)} \right\}$ using the Adam optimization during the interaction and the interaction force corresponding to all waypoints, i.e. $F_{h_t} = \left\{ F_{h_1}, \cdots, F^*_{h_i}, F^*_{h_{i+1}}, F^*_{h_{i+2}}, \cdots, F_{h_N} \right\}$. The idea of the proposed ILC algorithm is to update each waypoint such that the

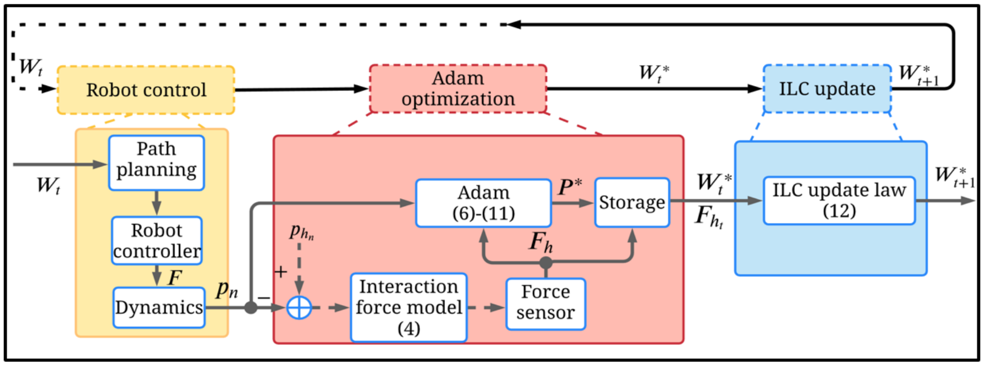

Fig. 3: The proposed control framework, including three units of robot control, Adam optimization and ILC update. The "storage" block is used to store the information e.g., waypoints and interaction force. The interaction force model is only for analysis purpose but not used to generate the interaction force. In practice, the interaction force is applied by the human user and measured by a force/torque sensor.

robot's actual path point gets close to the human partner's desired one. This is equivalent to making $F_h = \mathbf{0}$, so the ILC updating law is designed as

$$W_{t+1}^* = W_t^* + \lambda F_{h_t} \tag{12}$$

where $\lambda > 0$ is the learning rate. Then we set the robot's reference waypoints for the next iteration as $W_{t+1} = W_{t+1}^*$ and use $\left\{P_{(1,t+1)}, \cdots, P_{(i,t+1)}, \cdots, P_{(N,t+1)}\right\}$ to represent W_{t+1}.

In the following, we briefly analyze the convergence of the above ILC updating law. From Eq. (4), it can be obtained that

$$F_{h_t} = K_h(W_{tar} - W_t^*) \tag{13}$$

where W_{tar} represents the waypoint on the human partner's desired path. Combining Eqs. (12) and (13), we have

$$W_{t+1}^* - W_{tar} = W_t^* - W_{tar} + \lambda K_h(W_{tar} - W_t^*) \tag{14}$$

By defining $E_t^* = W_t^* - W_{tar}$, Eq. (14) can be rewritten as

$$E_{t+1}^* = E_t^* - \lambda K_h E_t^* \tag{15}$$

To prove the convergence of the updating law, we define a Lyapunov function candidate as below:

$$V_{(E_t^*)} = E_t^{*T} \mathbf{P} E_t^* \tag{16}$$

where $\mathbf{P}$ is a positive definite real symmetric matrix. The difference of the Lyapunov function candidate between two successive iterations is

$$\begin{aligned} \Delta V_{E_t^*} &= V_{E_{t+1}^*} - V_{E_t^*} \\ &= (E_t^* - \hat{K}_h E_t^*)^T \mathbf{P} (E_t^* - \hat{K}_h E_t^*) - E_t^{*T} \mathbf{P} E_t^* \\ &= E_t^{*T} \mathbf{P} E_t^* - \hat{K}_h E_t^{*T} \mathbf{P} E_t^* - E_t^{*T} \mathbf{P} \hat{K}_h E_t^* \\ &\quad - \hat{K}_h E_t^{*T} \mathbf{P} \hat{K}_h E_t^* - E_t^{*T} \mathbf{P} E_t^* \\ &= \hat{K}(\hat{K}_h - 2) E_t^{*T} \mathbf{P} E_t^* \end{aligned} \tag{17}$$

where $\hat{K}_h = \lambda K_h$. Since λ can be set as small enough such that $0 < \hat{K}_h < 2$, it yields $\Delta V_{E_t^*} \leq 0$. Moreover, $\Delta V_{E_t^*} = 0$ only when $E_t^* = 0$, otherwise $\Delta V_{E_t^*} < 0$ indicating that $V_{E_t^*}$ monotonically decreases. Therefore, we can conclude that

- $\lim_{t \to \infty} E_t^* = 0$ which indicates that $\lim_{t \to \infty} W_t^* = W_{tar}$, hence the robot's updated waypoints using the ILC algorithm can track the waypoints on the human partner's desired path.
- when $\lim_{t \to \infty} W_t^* = W_{tar}$, $F_{h_t} = \mathbf{0}$, which indicates that the robot's waypoints will not be updated and the convergence is achieved.

The details of the Adam optimization and the ILC updating algorithm are summarized in **Algorithm 1**.

IV. Simulation

In this section, simulation results are presented to demonstrate the advantages of the proposed waypoints updating method by comparing with the existing ones: Adam optimization in [17] and ILC in [14]. The aforementioned HRI scenario is considered, where a human user guides a robotic manipulator to complete a path following task. The human user's desired path is unknown to the robot and needs to be estimated by using the information of interaction force. The control objective is to make the robot's actual path follow the human user's desired one.

Considering the waypoints updating method proposed in the paper, enough waypoints are selected to represent the human's desired path. For the motion planning between two waypoints, we set a constant speed. The specific parameters in the simulation are given in Table I, and the number of the waypoints N is set to 31, the learning rate λ in the ILC updating law and the stiffness parameter K_h are set to satisfy the convergence of the proposed ILC. The other parameters are selected to be the same as the ones given in [16].

The human user's desired path is set as a sinusoidal path, in which the initial waypoint P_0 and the last waypoint P_N are set as $(-0.7, 0.44, 0)m$ and $(-0.1, 0.44, 0)m$, respectively. The corresponding waypoints are designed as $W_{tar} = \{-0.7+0.2i,\ 0.2sin(\frac{10}{3}\pi(-0.3+0.2i))+0.44,\ 0\}m$ where $i = 0, 1, 2, 3 \cdots$ and $i \in [0, 30]$.

A. Convergence speed

Fig. 4(a) and Fig. 4(b) show the path updating in seven iterations by the Adam optimization method [17] and spatial

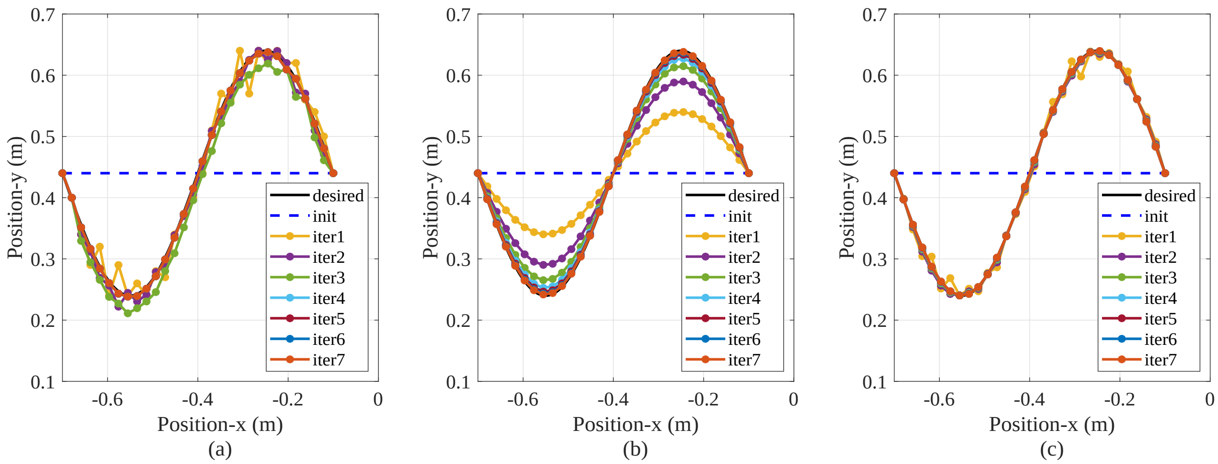

Fig. 4: Sinusoidal path following by the proposed method and two existing methods: (a) Adam path optimization algorithm; (b) ILC updating method; (c) Proposed method.

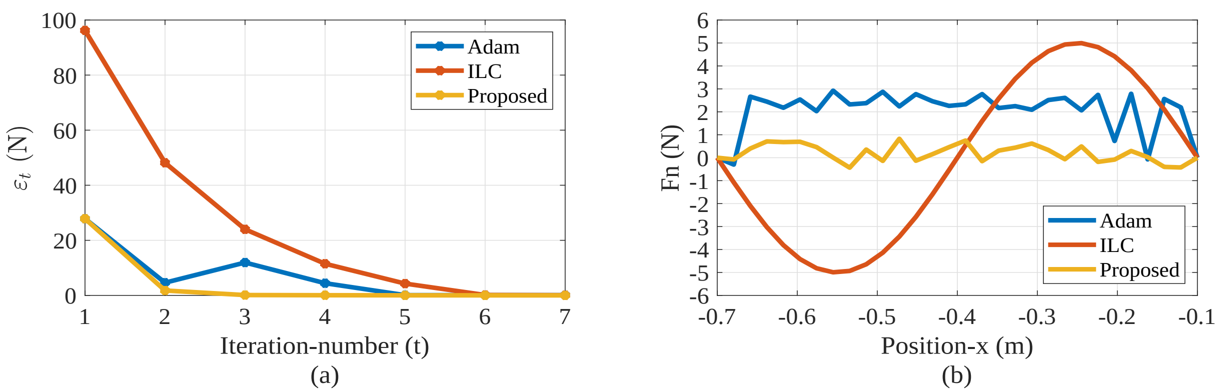

Fig. 5: (a) The force function ε_t of the three methods varies with the number of iterations. (b) The interaction force of the three methods in the third iteration.

ILC method [14], respectively. It is found that the Adam optimization method needs five iterations and the spatial ILC method needs seven iterations for following the desired path. Compared with these two methods, the proposed one only needs three iterations to achieve similar path following performance. These results clearly show that the proposed method achieves faster learning.

B. Force function and interaction force

In order to compare the aforementioned three methods from the perspective of human effort reduction, the force function values defined in Eq. (6) under three methods are shown in Fig. 5(a). Since the number of waypoints N is fixed and the path points are evenly distributed along the X axis, the force function ε_t can effectively reflect the effort made by the human user during the interaction. It can be found that the proposed method needs only three iterations to reduce the ε_t close to zero. In comparison, the other two methods need more iterations so the interaction requires the human user to make more effort.

Fig. 5(b) shows the interaction forces under three methods in the third iteration. In an ideal case, the interaction force reduced to zero means that the human partner does not need to guide the robot and correct the actual path, as the robot learns the desired path of the human user. In the third iteration, the interaction force under the proposed method reduces to less to 1N while larger forces are found under the other two methods.

The method proposed in this paper combines the advantage of the Adam algorithm in [17] for real-time path optimization and the advantage of the ILC algorithm to improve the learning performance iteratively [14]. The above simulation results verify the superiority of the proposed method in terms of faster learning and reduced human effort, by comparing with the Adam algorithm and spatial ILC method.

V. Experiment

In this section, the validity of the proposed method is further verified by experiments on a robotic platform Sawyer, which has been developed by Rethink Robotics [20]. The experimental scenario is shown in Fig. 6, where the Sawyer robot is guided by a human hand along a given path on a whiteboard (see the attached video). The interaction force between the human hand and the robot is measured by the force/torque sensor Robotiq FT300 at the end-effector of the robot, and the communication interface in the robot operating system (ROS) is used to collect the interaction force data. The

Algorithm 1: Adam+ILC algorithm for updating waypoints

Input: Initial waypoints $\mathbf{W_0}$;
Output: Updated waypoints $\mathbf{W_{t+1}}$;
Initialisation parameters: m_t, v_t, β_1, β_2, λ, $t = 0$, N, R, t_{max} is the maximum number of iterations.
while $t < t_{max} \cup \varepsilon_t > 0$ **do**
  $t+ = 1$
  $\varepsilon_t = 0$
  for $i \in [1, N)$ **do**
    $Updatewaypoints(P_i, P_{i+1}, m_{t-1}, v_{t-1}, t, W_t, \varepsilon_t)$
    $m_t \leftarrow$ average of all m_t
    $v_t \leftarrow$ average of all v_t
  end
  $W_{t+1} = W_t^* + \lambda F_{h_t}$
end

Function
$Updatewaypoints(P_{start}, P_{end}, m_{t-1}, v_{t-1}, t, W_t, \varepsilon_t)$
Generate path points p_n from P_{start} to P_{end} using (Sawyer's) motion planning library.
for *each path point* p_n **do**
  if $\| F_{h_n} \| > f_\tau$ **then**
    $m_t = \beta_1 m_{t-1} + (1 - \beta_1)\mathbf{E}F_h$
    $v_t = \beta_2 v_{t-1} + (1 - \beta_2)\mathbf{E}F_h^2$
    $\hat{m}_t = \frac{m_t}{1-\beta_1^t}$
    $\hat{v}_t = \frac{v_t}{1-\beta_2^t}$
    $P^* = [P_{end_x} + \frac{\alpha \cdot \hat{m}_{t_x}}{\sqrt{\hat{v}_{t_x}+\epsilon}}, P_{end_y} + \frac{\alpha \cdot \hat{m}_{t_y}}{\sqrt{\hat{v}_{t_y}+\epsilon}}, P_{end_z} + \frac{\alpha \cdot \hat{m}_{t_z}}{\sqrt{\hat{v}_{t_z}+\epsilon}}]$
    $\varepsilon_t + = F_{h_n}^*$
    if $|\overrightarrow{p_n P^*}| > R$ **then**
      $Updatewaypoints(p_n, P^*, m_{t-1}, v_{t-1}, t, W_t, \varepsilon_t)$
    else
      $P_{end} = P^*$
    end
  end
end
end

TABLE I: Parameters Setting

β_1	β_2	ϵ	α	λ	K_h	R	f_τ
0.9	0.999	10^{-8}	0.01	0.015	$10N/m$	$0.002m$	$1N$

waypoints are updated based on the interaction force and the position information, and then they are sent to the Sawyer robot for execution.

As is shown in Fig. 6, the movement of the end-effector is constrained in the XOY plane. The desired path is defined as $y_d = (0.18sin\frac{x_d}{0.28}\pi + 0.5)m, x_d \in [0, 0.56]$, which is unknown to the robot, so the initial waypoint P_0 is set to $(0, 0.5, 0)m$, and the last waypoint P_N is set to $(0.56, 0.5, 0)m$ in the coordinate frame of the Sawyer. The number of the waypoints N is set to 41, and the initial waypoints

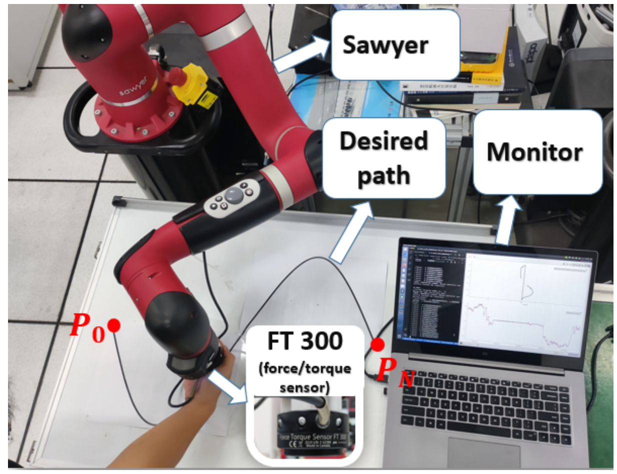

Fig. 6: Experimental scenario: the Sawyer robot is guided by a human hand along a given path on a whiteboard.

$W_0 = \{(0, 0.5, 0), (0.014, 0.5, 0), \cdots, (0.56, 0.5, 0)\}\, m$ are uniformly distributed on the path from the start point P_0 to the end point P_N. The threshold f_τ is set to $2N$. The learning rate α of Adam is set to 0.015 and the learning rate λ of ILC is set to 0.005. The other parameters are the same as set in the simulations, given in Table I.

A. Path following

Fig. 7(a) shows the actual path of the robotic arm's end-effector under the proposed method in the experiment. Each reference path in each iteration is composed of the updated waypoints and the linear point-to-point motion planning strategy of the Sawyer robot is adopted. Within an iteration, all the waypoints are updated online according to the Adam optimisation. At the end of each iteration, all the waypoints are further updated according to the ILC method and they are used to form the reference path of the robot in the next iteration. Through the observation of Fig. 7(a), we can find that the actual path gradually gets close to the desired path and it converges after 3 iterations. Correspondingly, the interaction force is significantly reduced as the iteration number increases and converges to a range of values close to zero, as illustrated in Fig. 8(a). As the actual path of the robot approximately coincides with the desired path, the human user only needs to apply a small force to correct the actual path when it deviates from the desired one, so human load is significantly reduced. When there is no interaction force applied, the robotic arm actively moves along the desired path, indicating transfer of the task knowledge from the human user to the robot.

To illustrate the advantages of the proposed method, the same task is carried out by using Sawyer's impedance control and setting its Z-axis stiffness with a maximum value and X and Y axes as 0. As shown in Fig. 7(b), the human user can move the end-effector to follow the desired path, but it is difficult to ensure accurate following, especially at the large-curvature areas on the path. This corresponds to large interaction forces as shown in Fig. 8(b). In particular, the interaction force does not decrease with the iteration number

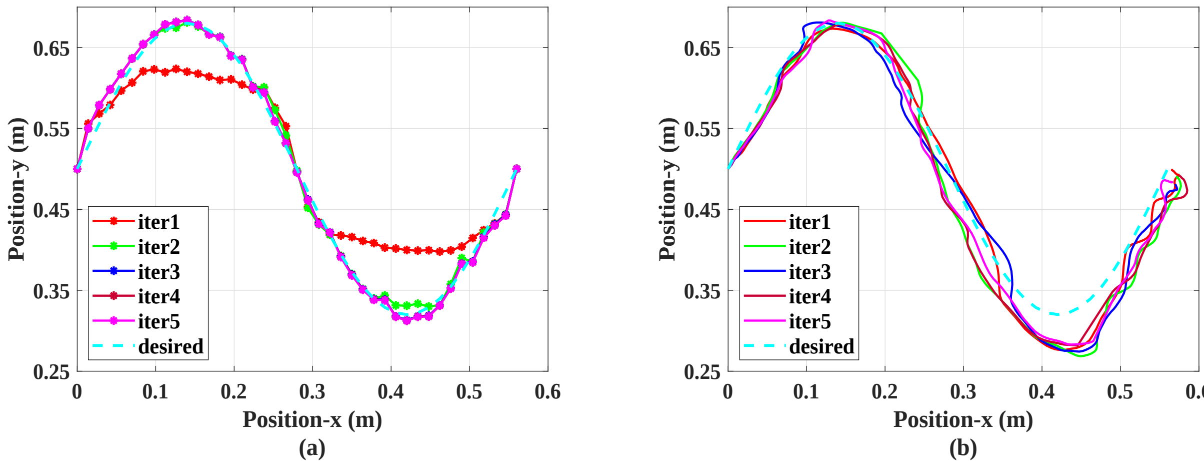

Fig. 7: (a) The actual path under the proposed method. (b) The actual path under impedance control.

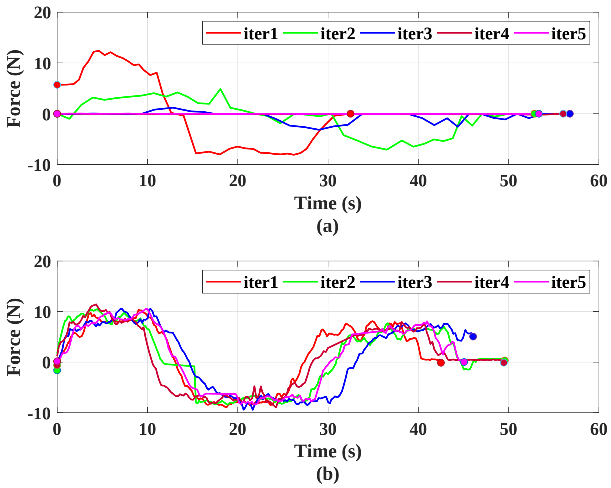

Fig. 8: (a) The interaction force under the proposed method. (b) The interaction force under impedance control.

without the learning capability, and it is significantly larger than that under the proposed method, indicating requirement of larger human effort.

B. Robustness against human's movement uncertainty

Due to the human's movement uncertainty, the interaction period of each iteration is different as shown in Table II and Fig. 8(a). During the first iteration, the human user plays a leader role and applies relatively larger forces throughout the entire interaction process. Therefore, the optimization is triggered to add new waypoints based on the Adam optimisation algorithm. With the more intermediate waypoints updated, the second iteration takes a longer time period. As the interaction force is significantly reduced from the second iteration, the optimization is less triggered so the time period of each iteration converges to a range of similar values. However, as the movement of the robot is affected by the human's applied force, there is still a small variation in the iteration period. Despite the uncertainty, the path learning method based on waypoints proposed in this paper can effectively update the

TABLE II: Iteration Periods (s)

$T1$	$T2$	$T3$	$T4$	$T5$
32.49	52.85	56.77	56.04	53.34

path to reduce the interaction force, showing its robustness. This is a favorable property that traditional ILC does not have, which has a strong assumption of task repetitiveness with either a temporal or a spatial period, thus not suitable for HRI.

In order to verify the robustness of the proposed method against uncertainties due to different human users, three male human subjects aged 21-26 are recruited in the pilot experiments. They have prior knowledge about robotics but do not know the proposed method in this paper. They are instructed to hold the robot's end-effector and move the robot along the given path, at a speed they feel comfortable with. Each human subject performs 5 iterations in one task and repeat this task for 5 times.

The average path error, the maximum interaction force, and the force function value during each iteration of the learning process are shown in Fig. 9. The average path errors of the three users shown in Fig. 9(a) can all be kept within $0.015m$ after 3 iterations. When the error between the actual path of the end-effector and the desired path of the human user is too small to be identified by the human user, there is no need to apply a force to correct the path of the robot arm so the interaction force F_h will converge to a value near 0 (a small non-zero value due to the measurement noise of the force sensor), as shown in Fig. 9(b). Correspondingly, Fig. 9(c) shows the force function ε_t converging to 0 after several iterations. Despite small differences between results of different human subjects, coherent performance of the proposed method is achieved.

VI. Conclusions

In this paper, a method for learning of the desired path of the human partner is proposed for human-robot interaction (HRI). This method combines the advantages of the Adam optimization algorithm and the iterative learning control (ILC).

From the results of the simulations, it can be confirmed that the proposed method can effectively learn the human partner's desired path that is unknown to the robot. Compared with the existing methods, the proposed one increases the learning speed while reducing the human effort. Through further comparative experiments, the feasibility of the proposed method and its robustness against human uncertainties are verified.

This paper considers a general scenario of HRI, and the proposed approach needs to be customized when it is applied to a specific application. For example, when it is used in load transportation, problems such as the relative motion between the robot and the load, manipulation of a long object requiring the human user and the robot to hold its two sides need to be addressed. When applying the proposed method to a 3D case, the coupling between different directions may have a more significant effect compared to a 2D case, e.g. small tracking errors in three directions may lead to a large error. Moreover, the proposed method mainly deals with the interaction between the human and the robot, but does not consider the interaction between the robot and the environment where force control needs to be considered. Also, the selection of the robot's start and end positions on its reference path is not studied in this paper. All these limitations and/or problems need to be further investigated.

Finally, relevant to the proposed method is the machine learning method extensively studied for trajectory learning in HRI. Actually, it is possible to combine machine learning methods and ours, e.g. using machine learning methods to learn an initial trajectory and our method to modify it online; or using machine learning methods for high-level decision making and our method for low-level motion planning and control.This is another interesting topic worth exploring in future works.

REFERENCES

[1] S. Grahn, B. Langbeck, K. Johansen, and B. Backman, "Potential advantages using large anthropomorphic robots in human-robot collaborative, hand guided assembly," *Procedia CIRP*, vol. 44, pp. 281–286, 2016.

[2] P. Salvine, M. Nicolescu, and H. Ishiguro, "Benefits of human-robot interaction," *IEEE Robotics and Amp Amp Automation Magazine*, vol. 18, no. 4, pp. 98–99, 2011.

[3] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Survey: Robot programming by demonstration," *handbook of robotics*, 2008.

[4] T. Ren, Y. Dong, D. Wu, and K. Chen, "Design of direct teaching behavior of collaborative robot based on force interaction," *Journal of Intelligent & Robotic Systems*, vol. 96, no. 1, pp. 83–93, 2019.

[5] C. H. Park, J. H. Kyung, D. I. Park, K. T. Park, D. H. Kim, and D. G. Gweon, "Direct teaching algorithm for a manipulator in a constraint condition using the teaching force shaping method," *Advanced Robotics*, vol. 24, no. 8, pp. 1365–1384, 2010.

[6] D. Wang and C. C. Cheah, "An iterative learning-control scheme for impedance control of robotic manipulators," *The International Journal of Robotics Research*, vol. 17, no. 10, pp. 1091–1104, 1998.

[7] A. Tayebi, "Adaptive iterative learning control for robot manipulators," *Automatica*, vol. 40, no. 7, pp. 1195–1203, 2004.

[8] P. M. Aubin, M. S. Cowley, and W. R. Ledoux, "Gait simulation via a 6-dof parallel robot with iterative learning control." *IEEE Trans Biomed Eng*, vol. 55, no. 3, pp. 1237–1240, 2008.

[9] Y. Li and S. S. Ge, "Impedance learning for robots interacting with unknown environments," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1422–1432, 2014.

[10] X. X. He, Z. H. Qin, and D. Zhang, "Adaptive iterative learning control of robot manipulators in the presence of environmental constraint," *Control Theory and Applications*, pp. 133–150, 2012.

[11] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Systems*, vol. 26, no. 3, pp. 96–114, 2006.

[12] X. Li, J.-X. Xu, and D. Huang, "An iterative learning control approach for linear systems with randomly varying trial lengths," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1954–1960, 2013.

[13] X. Li, J.-X. X, and D. Huang, "Iterative learning control for nonlinear dynamic systems with randomly varying trial lengths," *International Journal of Adaptive Control and Signal Processing*, vol. 29, no. 11, pp. 1341–1353, 2015.

[14] J. Xia, Y. Li, L. Yang, and D. Huang, "Spatial repetitive learning control for trajectory learning in human-robot collaboration," *IEEE 58th Conference on Decision and Control (CDC)*, pp. 5568–5573, 2019.

[15] J. Xia, D. Huang, Y. Li, and N. Qin, "Iterative learning of human partner's desired trajectory for proactive human-robot collaboration," *International Journal of Intelligent Robotics and Applications*, vol. 4, no. 2, pp. 229–242, 2020.

[16] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Computer Science*, 2014.

[17] Y. Gao, H. J. Chang, and Y. Demiris, "Iterative path optimisation for personalised dressing assistance using vision and force information," *IEEE/RSJ international conference on intelligent robots and systems (IROS)*, pp. 4398–4403, 2016.

[18] A. G. Feldman, "Once more on the equilibrium-point hypothesis (λ model) for motor control," *Journal of Motor Behavior*, vol. 18, no. 1, pp. 17–54, 1986.

[19] D. W. Franklin, E. Burdet, K. P. Tee, R. Osu, C. M. Chew, T. E. Milner, and M. Kawato, "Cns learns stable, accurate, and efficient movements using a simple algorithm," *The Journal of neuroscience*, vol. 28, no. 44, 2008.

[20] A. Smith, C. Yang, C. Li, H. Ma, and L. Zhao, "Development of a dynamics model for the baxter robot," *IEEE International Conference on Mechatronics and Automation*, pp. 1023–1028, 2016.

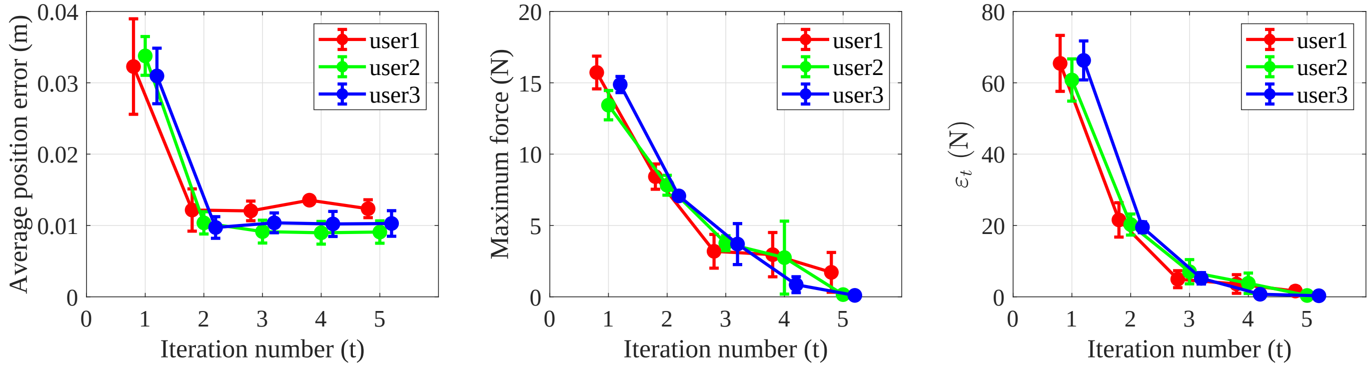

Fig. 9: (a) Average path error, (b) maximum interaction force and (c) force function value of each iteration for different users.