Human-Like Robot Action Policy Through Game-Theoretic Intent Inference for Human–Robot Collaboration

Yubo Sheng, *Graduate Student Member, IEEE*, Yiwei Wang, *Member, IEEE*, Haoyuan Cheng, Huan Zhao, *Member, IEEE*, and Han Ding, *Senior Member, IEEE*

***Abstract*—Harmonious human–robot collaboration requires the robot to behave like a human partner, which raises the critical question of what factors make the robot do so. This article proposes a series of policies based on empathetic and nonempathetic intent inference, proactive and reactive action planning, and ego and nonego action styles to examine, which modules enable robots to exhibit human-like behaviors. Two series of experiments are conducted with human subjects to test the performance of the proposed controllers. In Experiment 1, the participant must identify whether the collaborating partner is a human, similar to a turing test. The classification results empirically verify that the designed empathetic proactive policies enable the robot to exhibit human-like behaviors. Experiment 2 indicates that the proposed policy can be applied to complex collaborative tasks, and this result is consistent with the findings of Experiment 1. From empirical evidence from the experiments, we believe that empathy and proactive policies are essential elements to enable robots to perform human-like actions.**

***Index Terms*—Cognitive human–robot interaction, cooperating robots, human factors and human-in-the-loop, physical human–robot interaction.**

Nomenclature

RE	Reactive ego.
RNE	Reactive nonego.
PE	Proactive ego.
PNE	Proactive nonego.
HH	Human–human.
RCAH	Robot is classified as the human.
RCAR	Robot is classified as the robot.
HCAH	Human is classified as the human.
HCAR	Human is classified as the robot.

Received 1 March 2025; accepted 4 August 2025. Date of publication 27 August 2025; date of current version 22 September 2025. This work was supported in part by the National Science Foundation of China under Grant 52188102, Grant 62203180, and Grant 52090054, in part by the Hubei Science and Technology Major Program under Grant 2023BCA002 and Grant 2023BAA016, and in part by the Taihu Lake lnnovation Fund for Future Technology, HUST: 2023-A-2. This article was recommended for publication by Associate Editor Dongheui Lee and Editor Dana Kulic upon evaluation of the reviewers' comments. *(Corresponding author: Yiwei Wang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Institutional Review Board (IRB) of the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology under Application No. IORG0003571.

Yubo Sheng, Haoyuan Cheng, Huan Zhao, and Han Ding are with the State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: sheng_yubo@hust.edu.cn; m202170692@hust.edu.cn; huanzhao@hust.edu.cn; dinghan@hust.edu.cn).

Yiwei Wang is with the State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China, and also with the Institute of Medical Equipment Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: wang_yiwei@hust.edu.cn).

Digital Object Identifier 10.1109/TRO.2025.3603556

I. Introduction

ROBOTICS has been widely applied in various fields such as medical [1], industrial [2], and intelligent vehicles sectors [3] in recent years. Since fully autonomous robots remain impractical in many scenarios, human–robot collaboration (HRC), which combines the capabilities of both humans and robots, is essential [4]. HRC requires the robot and the human agents to collaborate as a team toward a common goal [5]. Research suggests that the performance of the collaboration improves when robots exhibit human-like behaviors [6], [7]. Therefore, many researchers devote efforts to making the robot behave like a human [8], [9] to achieve seamless HRC.

Research on human–human collaboration (HHC) suggests that the ability to understand the intent of the partner is a basic component of successful collaboration [10]. Hence, it is critical for the robot agent to incorporate the human's intent, which is inferred from prior observations, into planning its future actions properly [11]. Intent in collaborative robotics usually includes the goal of the agent or the latent parameter, which can be encoded as preferences that induce different action trajectories [12], [13], [14]. The machine theory of mind (ToM) discusses the capability of the agent to infer the other's intent [15]. First-order machine ToM enables the robot to infer human intent based on observations, while second-order machine ToM allows the robot to take the human's perspective and infer the human's understanding of the robot's intent [16]. However, intent inference is challenging because collaborating agents are generally unaware of each other's intent, making it a double-blind problem.

The study on a distributed multiagent collaborating system with double-blind intents indicates that adopting a second-order machine ToM mechanism, which allows all agents to take into account others' potential misunderstanding of their intents during intent inference, can significantly improve the convergence

of parameter estimation, while the first-order machine ToM inference mechanism would cause the system unstable [13]. Therefore, the robot needs to empathize with the human's perspective: they are unaware of the true intent of the robot and are trying to infer it. By adopting this perspective to infer intent, the robot becomes a more sophisticated and proactive system [17].

This article aims to design a control policy that enables the robot to perform human-like actions in real-time in an HRC scenario with double-blind intents and achieve effective collaboration. In this scenario, the goal of the task is clearly defined, but the agents are unaware of each other's action preferences. Therefore, agents need to infer the preferences of collaborating partners so that they can adapt to each other's actions for task completion. Psychological research has shown that human action tendencies in most scenarios can be categorized into two types: the approach tendency, which drives individuals to actively pursue a goal, and the avoidance tendency, which enables individuals to evade potentially negative outcomes [18]. Accordingly, this article translates this dual tendency into two action preferences to capture the agent's behavior in the HRC scenario: the active preference, which drives the agent to actively advance the task, and the cautious preference, which encourages the agent to avoid actions that could negatively impact collaboration.

This article presents an empathetic intent inference architecture that utilizes observed human actions to infer human intent based on the second-order machine ToM. The proposed method combines a game theory-based intent-action model with Bayesian prediction, enabling the robot agent to infer the human agent's intent by estimating the human's understanding of the robot's intent. As a result, the probability distributions of the human's intent and the human's expectation of the robot's intent can be obtained.

Based on the inference results, two robot controllers, referred to as reactive and proactive policies, are designed accordingly. The reactive policy directly utilizes the results of the inference to plan the robot's actions as similar to [16], while the proactive policy takes into account the human's dependency on the robot's next action when planning the robot's actions. In addition, we notice that few studies have discussed whether a robot should be an ego agent. Therefore, we design not only the ego robot employed in [17], but also a nonego robot for each strategy. The performance of the ego robot and the nonego robot is compared through experiments. In collaboration, the ego robot takes actions with a fixed preset intent, whereas the nonego robot adjusts its intent based on the inference outcome.

Two series of human subjects experiments are conducted to evaluate the feasibility of the proposed method. Quantitative metrics are defined to validate the effectiveness of the proposed policies. Several studies comprehensively assess the performance of the methods designed for the collaborative agent by investigating user evaluation through subjective metrics [19], [20], [21]. Therefore, subjective metrics are adopted to measure the subjective evaluation of the designed policies.

In Experiment 1, participants are required to identify whether their collaboration partners are humans or robots, similar to a turing test. The results of the partner classification empirically demonstrate that the proposed PNE method can make the robot's behavior align with the human behavioral patterns in the participants' cognition. The quantitative results indicate that the proposed PNE policy can achieve well-performed HRC. The subjective results show that the designed strategy can enable a good user evaluation. To provide some insights into HRC, synthesis metrics are derived from the classification results to identify the behavioral characteristics that could lead users to perceive their collaboration partner as human.

The role of an agent represents a behavioral pattern that naturally emerges during collaboration [21]. As suggested in [22], role can be categorized as *leader* and *follower*, where the *leader* puts more effort into performing the task than the one as a *follower*. Inspired by prior work, Experiment 1 employs role as a metric to investigate the behavioral patterns of agents under different collaborative conditions. We define the role index based on the recorded actions of agents and analyze the role exchanges during the collaborative process.

An important application scenario of collaborative robots is robot-assisted surgeries, which aim to enhance surgical precision and dexterity and reduce surgeons' workload [23], [24]. For a surgical robot, recognizing the surgeon's intent to execute the correct actions is critical for ensuring the success of the surgery [11]. Soft tissue manipulation is a common yet challenging task in surgical procedures [24]. Therefore, Experiment 2 investigates whether the robot can collaborate with the operator using the proposed policy to improve tissue manipulation performance while ensuring a good subjective evaluation. A virtual environment is designed to simulate the surgical scenario for this purpose. The experimental results demonstrate that the proposed policy improves the performance of goal tracking in the indirect simultaneous positioning task for soft tissue. A combined analysis of quantitative and subjective evaluations indicates that the proposed PNE strategy enables a more accurate, effortless, and satisfying collaboration than the nonempathetic ego baseline strategy. These findings validate the potential of the proposed method for application in a complicated collaborative task.

The main contributions of this article can be summarized as follows.

1) An empathetic intent inference framework is developed using a game theory-based human–robot collaborative model to address the double-blind issue.
2) A PNE action policy is proposed to implement human-like actions, validated through human subjects experiments.
3) Key characteristics that could contribute to the robot agent being subjectively perceived as human are identified through synthesized results.
4) The designed PNE policy is shown to enable effective HRC and handle a complex collaborative task.

The rest of this article is organized as follows. Section II provides an overview of the related works. Section III introduces the empathetic intent inference method and the robot action planning policies in the discussed scenario. Section IV details the human subjects experiment setup. Section V presents

the experimental results, and discusses the findings. Finally, Section VI concludes this article.

II. Related Work

A. Intent Inference in HRC

Diverse methods based on the first-order ToM have been applied to the intent inference in HRC. Various bioelectric signals generated by the human body, such as electroencephalogram and electromyogram, have been widely used for intent prediction in HRC [25], [26]. These acquired signals can be classified into different human action preferences using signal processing methods such as support vector machine [27] and neural network [28]. Observed human actions also serve as implicit indicators of human intents [29]. Hidden Markov model and partially observable Markov decision process model are often applied to model human intent based on past actions of both agents [11], [12]. Predicted human intent can be obtained by calculating the Markov model's transfer and observation probability matrices. Similarly, Bayesian prediction has been widely used in tasks where observed human actions are utilized to infer human intent [29], [30], [31]. Moreover, the data-driven intent and action prediction approaches, e.g., based on deep neural networks [32], [33] also provide promising solutions. Besides, game-theoretic models have garnered much interest in describing interactive behaviors between robot and human agents in collaborative tasks with incomplete information [34], [35]. In game theory, the high-level intents and low-level actions can be formulated as the cost functions of the agents [13]. Pezeshki et al. [36] modeled the assist-as-needed robotic rehabilitation as a nonzero-sum game.

Recent works begin to design more sophisticated intent inference frameworks based on the second-order ToM. The "blame all" strategy presented in [13] has been proven to make the multiagent system stable. However, the constraints set during the proof limit the complexity of the system. In a two-agent interaction scenario, Sun et al. [37] considered the impact of the robot's future actions in inferring human actions. However, the inference was based on the assumption that the human was clear about what actions the robot would take. The game-theoretic method in [17] enabled a courteous driving policy during vehicle interactions through an empathetic intent inference algorithm where the intents of both agents were jointly inferred. However, this method was only validated by simulating the human as a rational agent.

B. Control Policy Enabling Human-Like Behaviors

Recent work has incorporated human-like elements into the robot's body motion planning to improve the recognizability of robot behavior [38], [39]. A common technique relies upon empirical criteria extracted from the recorded human motion data. For instance, Lauretti et al. [40] used joint characteristics and human pose-to-target relationships in Cartesian space to compute a compatible desired trajectory for the robot. Alternatively, based on biomarkers and physical laws extracted from captured human motion, such as elbow elevation [41] and wrist position [42], biomimetics has been used to enable the robot to mimic the arm trajectories. To get rid of the dependence on large quantities of data, Micheal et al. [8] proposed an algorithm to automatically synthesize human-like variants of input motions, improving the clarity of the social robot's motions.

The goal of this article is to enable the robot to exhibit human-like behavioral characteristics in joint actions, even when a human collaborator has difficulty observing the robot's body motion. For example, in collaborative transport of a large, elongated object to a desired position, the robot's body motion may be obscured. Nevertheless, the robot can still enhance the human's understanding of its actions through human-like behavioral characteristics, such as the coordination ability to keep the ends of the object aligned and the magnitude of the force exerted on the object, to achieve successful collaboration. Therefore, we focus mainly on the robot's action decision-making policy rather than motion planning. Behavior cloning based on imitation learning is a promising approach for replicating human-like behavior [43], [44]. However, learning-based decision-making methods remain constrained by dataset quality and require further improvement [9], [44].

Decision-making algorithms that generate human-like actions take into account human action characteristics and styles [45], [46]. Jiang et al. [47] incorporated the psychological factors influencing human decision-making to enable the robot to make decisions in a human-like manner. In [9], human preferences for safety, comfort, and efficiency were incorporated in designing the robot's action strategy. Compared to these studies, this article empirically validates that the proposed policy can lead users to perceive the robot's behavior as human-like through human subjects experiments. In addition, we identify key characteristics of human-like behaviors that could align with user knowledge.

III. Methodology

This section presents an approach based on a game-theoretic model, enabling the robot to infer human intent for the scenario in which two agents have to accomplish a task by collaborating. Then, the robot action controllers are designed using the inference results. The robot's decision-making process is depicted in Fig. 1.

A. Task Scenario

Fig. 2 shows the task scenario where a human and a robot try to achieve a common goal x_{goal} in a task space (e.g., moving an object) with state $x \in \mathbb{R}^m$. The dynamics of the system in the discrete-time domain are presented as

$$x(k+1) = f(x(k), u_h(k), u_r(k)) \tag{1}$$

where $x(k+1) \in \mathbb{R}^m$ is the system state at time step $k+1$. Both agents can observe the state and apply different inputs to the object. Equation (1) indicates that the inputs of both agents jointly determine the state of the object, necessitating collaboration to accomplish the task.

Let the human and the robot be denoted as agent h and agent r, respectively. The candidate action $u_i \in \mathbb{R}^l$ of agent

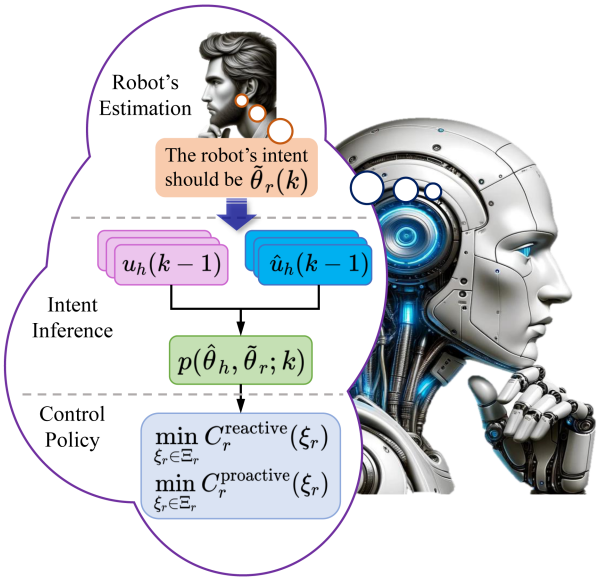

Fig. 1. Diagram illustrates the decision-making process. The robot infers the human's intent empathetically according to the estimation of the human's understanding of the robot's intent. The robot takes actions according to the inference results.

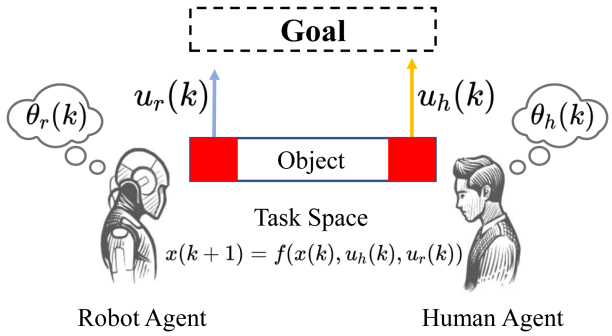

Fig. 2. Human–robot collaborative task is set up.

$i \in \{h, r\}$ is selected from a finite set U_i, which is specified as $U_i = \left\{u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(n)}\right\}$.

B. Model of HRC: A Game-Theoretic Approach Based on Nash Equilibrium

The collaboration model is formulated as a noncooperative two-player nonzero-sum game. When both agents know each other's intent, their optimal actions based on the Nash equilibrium are computed.

1) Cost Functions: The cost function of agent i should consist of two parts: the cumulative sum of the payoff of approaching the goal and the cost of actions, which is expressed in a finite horizon quadratic form as follows [48]:

$$C_h = \sum_{n=0}^{H-1} (\tilde{x}^T(k+n) Q_h \tilde{x}(k+n) + \theta_h u_h^T(k+n) u_h(k+n)) \tag{2}$$

$$C_r = \sum_{n=0}^{H-1} (\tilde{x}^T(k+n) Q_r \tilde{x}(k+n) + \theta_r u_r^T(k+n) u_r(k+n)) \tag{3}$$

where $\tilde{x}(k) = x(k) - x_{\text{goal}}$ is the error between the current and target states of the object. $H \in \mathbb{N}^+$ represents the control horizon of the system. $Q_i \in \mathbb{R}^{m \times m}$ is a constant weighting matrix, which penalizes the error between the current and target states of the object. $\theta_i \in \mathbb{R}^+$ is an intent parameter, which represents the preference in executing the action for each agent. The sequence of u_i in the control horizon is defined as the action trajectory $\xi_i(k) = \{u_i(k), u_i(k+1), \ldots, u_i(k+H-1)\}$. When necessary, the dependence of the variables on time step k will be omitted for brevity.

2) Nash Equilibrium Sets: The set of Nash equilibrium solutions based on the joint intent (θ_h, θ_r) at time k is defined as $\mathcal{Q}(\theta_h, \theta_r, k) = \{(\xi_h^*, \xi_r^*)\}$. According to the definition in [49], each element of $\mathcal{Q}$ is a pair of action sequences from both agents, satisfying the following conditions:

$$\xi_h^*(k) = \underset{\xi_h(k) \in \Xi_h}{\arg\min} C_h(x(k), \xi_h(k), \xi_r^*(k), \theta_h) \tag{4}$$

$$\xi_r^*(k) = \underset{\xi_r(k) \in \Xi_r}{\arg\min} C_r(x(k), \xi_h^*(k), \xi_r(k), \theta_r) \tag{5}$$

where Ξ_i is the set of permissible action sequences for the two agents. The first element of each optimal trajectory $\xi_i^*(k)[0]$ is selected as the agent's current probable action. $\cdot[0]$ represents the first element of the set. When both agents have full knowledge of each other's intents, they take actions according to (4) and (5).

C. HRC With Unknown Intent

In practice, neither the robot nor the human knows the intent of the other, rendering the Nash equilibrium solution derived in the previous section inapplicable. Therefore, the agent needs to infer the partner's intent to take actions. From the perspective of the robot, the solutions in (4) and (5) can be rewritten as

$$\hat{\xi}_h^*(k) = \underset{\xi_h(k) \in \Xi_h}{\arg\min} C_h(x(k), \xi_h(k), \xi_r^*(k), \hat{\theta}_h) \tag{6}$$

$$\xi_r^*(k) = \underset{\xi_r(k) \in \Xi_r}{\arg\min} C_r(x(k), \hat{\xi}_h^*(k), \xi_r(k), \theta_r) \tag{7}$$

where $\hat{\theta}_h$ represents the robot's inference of human intent, and $\hat{\xi}_h^*(k)$ is the human's planned action inferred by the robot. From (6) and (7), it is worth noting that the robot needs to derive the Nash equilibrium set from the human agent's perspective, which means the robot should empathetically understand how the human perceives its intent in order to infer the human's intent. Since the human is unaware of the robot's intent and the planned action, θ_r and ξ_r are modified to $\tilde{\theta}_r$ and $\tilde{\xi}_r^*$, respectively, to represent the robot's empathetical inference of the human's inference of the robot's intent and the planned action. In addition, $\mathcal{Q}(\theta_h, \theta_r, k)$ is rewritten as $\mathcal{Q}(\hat{\theta}_h, \tilde{\theta}_r, k) = \{(\hat{\xi}_h^*, \tilde{\xi}_r^*)\}$ to denote the Nash equilibrium set after the robot takes the perspective of

the human. Hence, the modified Nash equilibrium satisfies

$$\hat{\xi}_h^*(k) = \underset{\xi_h(k)\in\Xi_h}{\arg\min} C_h(x(k), \xi_h(k), \tilde{\xi}_r^*(k), \hat{\theta}_h) \tag{8}$$

$$\tilde{\xi}_r^*(k) = \underset{\xi_r(k)\in\Xi_r}{\arg\min} C_r(x(k), \hat{\xi}_h^*(k), \xi_r(k), \tilde{\theta}_r). \tag{9}$$

To avoid an infinite iteration of the inference process, the robot operates under the following assumptions.

Assumption 1: The human selects actions uniformly from $\mathcal{Q}(\hat{\theta}_h, \tilde{\theta}_r, k)$. This implies that from the robot's perspective, the human's actions follow the probability function:

$$p(\hat{\xi}_h; \hat{\theta}_h, \tilde{\theta}_r, k) \propto \frac{1}{\left|\{\hat{\xi}_h\}\right|}$$
$$\text{s.t.} \quad \hat{\xi}_h \in (\hat{\xi}_h, \tilde{\xi}_r) \text{ and } (\hat{\xi}_h, \tilde{\xi}_r) \in \mathcal{Q}(\hat{\theta}_h, \tilde{\theta}_r, k) \tag{10}$$

where $|\cdot|$ is the cardinality of a set.

Assumption 2: The human also assumes that the robot takes actions according to the same policy in (10).

Assumption 3: During collaboration, the two agents infer each other's intents at time k by observing each other's actions at time $k-1$.

D. Inference of Intent and Action

As the following equation shows, the inference problem is solved by identifying combinations of human and robot intents such that the inferred action sequence of the human at time $k-1$, denoted as $\hat{\xi}_h(k-1)$, should align its first element (denoted as $\hat{u}_h(k-1)$) with the observed human actions at time $k-1$ (denoted as $u_h(k-1)$). In other words, the robot's inferred estimation of what the human could have done should correspond to the human's actual actions

$$\min_{(\hat{\theta}_h, \tilde{\theta}_r)} ||\hat{\mathbf{u}}_h(k-1) - \mathbf{u}_h(k-1)||_2^2$$
$$\text{s.t.} \quad \hat{u}_h = \hat{\xi}_h[0] \text{ and } (\hat{\xi}_h, \tilde{\xi}_r) \in \mathcal{Q}(\hat{\theta}_h, \tilde{\theta}_r, k-1). \tag{11}$$

As mentioned in Section I, the action preferences of both agents in the collaboration are categorized into two types, implying that the intent set Θ is finite. The joint intent space $\Theta \times \Theta$ is used to describe combinations of the human and robot's intents. Since both Θ and Ξ are finite, the combinations of inferred intents that satisfy (11) can be obtained by enumerating over $\Theta \times \Theta$. The result of the enumeration is denoted as a set $\mathcal{S}(k) = \{(\hat{\theta}_h^*, \tilde{\theta}_r^*)_n\}_{n=1}^N$. Since each combination of intents is independent of the others, equal probability mass $(1/N)$ is assigned to each element of $\mathcal{S}(k)$ as follows in order to quantify the uncertainty of the inference results and integrate the uncertainty into action planning

$$p(\hat{\theta}_h, \tilde{\theta}_r; k) \propto \begin{cases} \frac{1}{N}, & \text{if } (\hat{\theta}_h, \tilde{\theta}_r) \in \mathcal{S}(k) \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

Equation (12) shows the computation of the empirical joint distribution $p(\hat{\theta}_h, \tilde{\theta}_r; k)$ defined on $\Theta \times \Theta$. In addition, the marginal distributions $p(\hat{\theta}_h; k)$ and $p(\tilde{\theta}_r; k)$ defined on Θ can be obtained as follows:

$$p(\hat{\theta}_h; k) = \sum_{\tilde{\theta}_r \in \Theta} p(\hat{\theta}_h, \tilde{\theta}_r; k) \tag{13}$$

$$p(\tilde{\theta}_r; k) = \sum_{\hat{\theta}_h \in \Theta} p(\hat{\theta}_h, \tilde{\theta}_r; k). \tag{14}$$

Equation (13) is the distribution of the robot's inference of the human intent, while (14) represents the distribution of the robot's estimation of the intent that the human thinks the robot should have.

From (10), each combination of intents $(\hat{\theta}_h^*, \tilde{\theta}_r^*) \in \mathcal{S}(k)$ can derive a conditional distribution of human actions $p(\hat{\xi}_h; \hat{\theta}_h^*, \tilde{\theta}_r^*, k)$ from the Nash equilibrium solution set $\mathcal{Q}(\hat{\theta}_h^*, \tilde{\theta}_r^*, k)$. Then, the marginal distribution $p(\hat{\xi}_h; k)$ can be calculated based on $p(\hat{\xi}_h; \hat{\theta}_h^*, \tilde{\theta}_r^*, k)$ and $p(\hat{\theta}_h; k)$. Similarly, the marginal distribution $p(\tilde{\xi}_r; k)$ can be obtained. However, the $p(\tilde{\xi}_r; k)$ is not the probability according to which the robot will take actions, it only represents the robot's inference of what the human estimates the robot would do.

It is worth noting that the agent with different intents may take the same action $u_i(\tau)$ at time τ, so using previous observations enables a more rigorous inference about the intent. Equation (11) can be rewritten as the following:

$$\min_{\{\tilde{\theta}_r(\tau)\}_{\tau=1}^{k-1}, \hat{\theta}_h} \sum_{\tau=1}^{k-1} ||\hat{\mathbf{u}}_h(\tau) - \mathbf{u}_h(\tau)||_2^2$$
$$\text{s.t.} \quad \hat{u}_h = \hat{\xi}_h[0] \text{ and } (\hat{\xi}_h, \tilde{\xi}_r) \in \mathcal{Q}(\hat{\theta}_h, \tilde{\theta}_r(\tau), \tau). \tag{15}$$

As a matter of fact, the solution of (15), represented by $\bar{\mathcal{S}}(k)$, can be found recursively based on the solution of (11). Specifically, let an inferred intent $\hat{\theta}_h$ exist in both $\bar{\mathcal{S}}(k-1)$ and $\mathcal{S}(k)$. The solutions in $\bar{\mathcal{S}}(k-1)$ that contain $\hat{\theta}_h$ can be written as $(\mathbf{a}, \hat{\theta}_h)$, where $\mathbf{a} = [a_1, \ldots, a_{k-1}] \in \mathcal{A} \subset \Theta^{k-1}$ is a sequence of robot's intents. In the same way, the solutions in $\mathcal{S}(k)$ that contains $\hat{\theta}_h$ can be written as $(b, \hat{\theta}_h)$, where b is a robot's intent. Appending b to the array $\mathbf{a}$ yields $[\mathbf{a}, b]$. Then, $([\mathbf{a}, b], \hat{\theta}_h)$ is a solution of (15) for all $\mathbf{a} \in \mathcal{A}$ and $b \in \mathcal{B}$. Following this property, it can be obtained that $\bar{p}(\hat{\theta}_h; k) \propto \bar{p}(\hat{\theta}_h; k-1) p(\hat{\theta}_h; k)$. If the inferred intent $\hat{\theta}_h$ does not exist in $\bar{\mathcal{S}}(k-1)$ or $\mathcal{S}(k)$, it will not exist in $\bar{\mathcal{S}}(k)$ and its deduced probability mass in $\bar{p}(\hat{\theta}_h; k)$ will be zero. The update of $\bar{p}(\hat{\theta}_h; k)$ leads to the updates of the joint probability $p(\hat{\theta}_h, \tilde{\theta}_r; k)$ and the marginal probability $p(\hat{\xi}_h; k)$, which are represented as $\bar{p}(\hat{\theta}_h, \tilde{\theta}_r; k)$ and $\bar{p}(\hat{\xi}_h; k)$, respectively. Specifically, the updates of these distributions are as follows:

$$\bar{p}(\hat{\theta}_h, \tilde{\theta}_r; k) \propto p(\hat{\theta}_h, \tilde{\theta}_r; k) \bar{p}(\hat{\theta}_h; k)$$
$$\bar{p}(\hat{\xi}_h; k) = \sum_{(\hat{\theta}_h, \tilde{\theta}_r) \in \Theta \times \Theta} p(\hat{\xi}_h; \hat{\theta}_h, \tilde{\theta}_r, k) \bar{p}(\tilde{\theta}_r, \hat{\theta}_h; k). \tag{16}$$

E. Action Controller

This section introduces two robot action controllers that leverage the result of the intent inference from Section III-D, namely, the reactive and proactive controllers. In addition, each controller is further divided into an ego robot controller and

a nonego robot controller depending on whether the robot has a fixed intent or not. Under the ego control policy, the robot has a specific intent θ_r, and the loss incurred from taking an action depends solely on the robot itself. In contrast, when the nonego control strategy is adopted, the robot becomes more adaptable. Its intent varies according to its inference of the human's estimation of the robot's intent $\tilde{\theta}_r$, i.e., the robot takes into account the human's expectations when evaluating the loss from its action.

1) Reactive Controller: When the distribution of the human's future actions $\bar{p}(\hat{\xi}_h; \hat{\theta}_h, k)$ is obtained, a reactive robot plans its actions by minimizing the expected loss in a control horizon. Two reactive controllers are then given, depending on whether the robot is ego, as follows:

RE controller:

$$\min_{\xi_r \in \Xi_r} C_r^{\text{reactive,ego}}(\xi_r) := \mathbb{E}_{\hat{\xi}_h \sim \bar{p}(\hat{\xi}_h;k)} \left[C_r \left(\xi_r, \hat{\xi}_h, \theta_r \right) \right] \quad (17)$$

RNE controller:

$$\min_{\xi_r \in \Xi_r} C_r^{\text{reactive, nonego}}(\xi_r) := \mathbb{E}_{\hat{\xi}_h \sim \bar{p}(\hat{\xi}_h;k)} \left[\sum_{\tilde{\theta}_r \in \Theta} p(\tilde{\theta}_r) C_r \left(\xi_r, \hat{\xi}_h, \tilde{\theta}_r \right) \right]. \quad (18)$$

As Ξ_r is a finite set, the solutions of (17) and (18) can be found by enumeration. The same logic applies to a proactive robot.

2) Proactive Controller: A proactive robot agent takes an action by considering the dependency of the human's action on the robot's next action. In detail, the robot computes the conditional distribution $\bar{p}(\hat{\xi}_h; \xi_r, k)$ according to ξ_r and $\bar{p}(\hat{\theta}_h; k)$, with the assumption that the human will promptly respond to ξ_r. For this purpose, the robot is required to obtain the set of optimal human actions for each $\hat{\theta}_h \in \Theta$ when ξ_r is given. This set is represented by $\mathcal{Q}_h(\xi_r) = \cup_{\hat{\theta}_h \in \Theta} \mathcal{Q}_h(\xi_r, \hat{\theta}_h)$ where $\mathcal{Q}_h(\xi_r, \hat{\theta}_h) = \{\hat{\xi}_h^*; \hat{\xi}_h^* = \arg\min_{\xi_h \in \Xi_h} C(\xi_h, \xi_r, \hat{\theta}_h)\}$. Then, for every element $\hat{\xi}_h^* \in \mathcal{Q}_h(\xi_r)$, it can be obtained

$$\bar{p}(\hat{\xi}_h^*; \xi_r, k) = \sum_{\hat{\theta}_h \in \Theta} \frac{\bar{p}(\hat{\theta}_h; k) 1(\hat{\xi}_h^* \in \mathcal{Q}_h(\xi_r, \hat{\theta}_h))}{|\mathcal{Q}_h(\xi_r, \hat{\theta}_h)|} \quad (19)$$

where $1(\cdot)$ is an indicator function. For $\hat{\xi}_h \in \Xi_h / \mathcal{Q}_h(\xi_r)$, we set $\bar{p}(\hat{\xi}_h; \xi_r, k) = 0$.

Then two types of proactive robot controllers are formulated according to whether the robot is ego

PE controller:

$$\min_{\xi_r \in \Xi_r} C_r^{\text{proactive,ego}}(\xi_r) := \mathbb{E}_{\hat{\xi}_h \sim \bar{p}(\hat{\xi}_h;\xi_r,k)} \left[C_r \left(\xi_r, \hat{\xi}_h, \theta_r \right) \right] \quad (20)$$

PNE controller:

$$\min_{\xi_r \in \Xi_r} C_r^{\text{proactive, nonego}}(\xi_r) := \mathbb{E}_{\hat{\xi}_h \sim \bar{p}(\hat{\xi}_h;\xi_r,k)} \left[\sum_{\tilde{\theta}_r \in \Theta} p(\tilde{\theta}_r) C_r \left(\xi_r, \hat{\xi}_h, \tilde{\theta}_r \right) \right]. \quad (21)$$

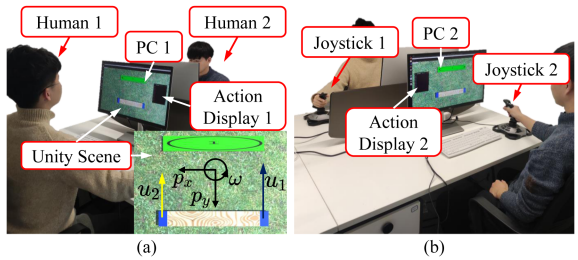

Fig. 3. Setup of Experiment 1. Human operators can push the object created by the UNITY 3-D physics engine with different forces through the joysticks. Two computers are used as intelligent robots. The operator needs to push the object to the target position through HRC or HHC. (a) PC1 side. (b) PC2 side.

IV. Experiment

In this section, two Institutional Review Board (IRB)-approved human subjects experiments are conducted to evaluate the performance of the designed robot action control policies in comparison to the HHC. Experiment 1 tests whether the proposed policy enables the robot to be subjectively recognized as a human through a collaborative scenario of moving an object to the target position. Experiment 2 further explores the potential applications by considering a surgical task that involves manipulating soft tissue to the desired location.

A. Participant Recruitment

A total of 28 participants (17 males and 11 females; age range: 20–34 years) from different backgrounds took part in Experiment 1; 16 participants (9 males and 7 females; age range: 23–35 years) were recruited for Experiment 2. Except for two left-handed participants in Experiment 2, all others were right-handed. None of the participants had prior experience with HRC or any disabilities. Furthermore, they were all unaware of the experiment's purpose beforehand. The study was supervised by the IRB of the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology with permission No. IORG0003571.

B. Experiment 1: Moving an Object

Experiment 1 follows a Turing test approach, where participants are tasked with identifying unknown collaborative partners. The classification results are used to determine whether the designed robot agent is perceived as exhibiting human-like action characteristics.

1) Hardware Setup: The framework of the Experiment 1 is shown in Fig. 3. The object moving task is developed using the UNITY 3-D Physics Engine. The top view of the scenario is shown to the user. Two Extreme 3-D Pro joysticks serve as human input devices, which are connected to two PCs through two USB ports. The black coordinates are the reference coordinates for the scenario. Each operator can only apply a force perpendicular to one end of the object by manipulating the joystick. The user needs to press the button on the joystick to push one end of the object with different forces. The force applied by the operator is displayed on the screen to facilitate

action adjustments. The two PCs used in the experiment are connected via a network cable for HHC, and each PC can also independently play the role of an intelligent robot to push the other end of the object in HRC. As shown in Fig. 3(a), the operator in PC 1 pushes the right end of the object, while in Fig. 3(b), the operator in PC 2 pushes the left. Both collaborative platforms are implemented using the robot operating system (ROS).

2) Parameter Setting: The mass of the object is set to $m = 0.1417$ kg. Its center of mass position is denoted by a vector $p = [p_x, p_y]^T$. The velocity at the center of mass is represented as $v = [v_x, v_y]^T$. The object's orientation and angular velocity are expressed as ω and $\dot{\omega}$, respectively. Then, the state vector $x \in \mathbb{R}^6$ of the system can be defined as $x = [p_x, p_y, v_x, v_y, \omega, \dot{\omega}]^T$. The initial state of the object is set as $x_0 = [0, 1.5, 0, 0, 0, 0]^T$, with a goal state of $x_{\text{goal}} = [0, 0, 0, 0, 0, 0]^T$. The finite set of candidate actions is defined as $U_i = \{-1.3, -1.0, -0.8, 0, 0.8, 1.0, 1.3\}$. The penalty matrix Q_i is denoted as $\text{diag}(30, 60, 1, 1, 3000, 1)$. The intent set, which describes the action preference of the agent, is set to $\Theta_i = \{0.001, 1\}$. When $\theta_i = 0.001$, the agent takes active actions, which means that the agent tends to use maximum force and actively advance the task during collaboration. Conversely, when $\theta_i = 1$, the agent acts cautiously, preferring minimal force or inaction to avoid excessive movement that could disrupt collaboration.

3) Experimental Tasks and Conditions: The main objective of this experiment is to verify whether the designed PNE action control policy enables the robot to behave like a human during HRC and achieve a good collaboration effect. Specifically, it is necessary to verify whether the robot's behavior under the PNE condition makes participants perceive their partner as human-like and whether collaboration under the PNE condition achieves better collaborative effectiveness compared to other collaboration conditions.

To fulfill this objective, participants are required to adhere to the following guidelines.

1) Push the object to the target position as accurately as possible while collaborating with the unknown partner.
2) Coordinate actions with the partner to keep the ends of the object aligned along the y-axis as much as possible while moving it.
3) Avoid moving the object backward, even if the object surpasses the target position.
4) Identify whether the collaborating partner is a human agent or a robot agent after each trial.
5) Avoid talking to each other during the whole experiment.

We acknowledge the inherent challenge in distinguishing between HRC and HHC in our experimental scenario. If a simple control policy can already convince participants that the robot behaves like a human, the classification results would lose significance. To address this concern, we introduce a baseline robot action control strategy to assess whether users can easily identify it as a robot. The baseline imitates a specific type of human behavior—applying greater force to the object when it is far from the target position and stopping when the object is very close to the target position. However, unlike humans, this baseline robot follows a fixed action policy solely based on the state of the object, without considering the action or intent of the human agent.

TABLE I
QUESTIONNAIRE 1 OF THE HUMAN SUBJECTS EXPERIMENTS FOR SUBJECTIVE METRICS

	Question
Partner-rating	Please assess your partner's performance.
Self-rating	Please assess your performance.
Overall-rating	Please assess the goal-tracking error achieved in this collaboration.
Classification	Do you think your collaborating partner is a human or a robot?

Participants collaborate with five different robot agents, each employing a distinct control strategy: RE, PE, RNE, PNE, and a baseline control strategy. In addition, participants are required to complete the task under an HH experimental condition. The ego robot agents are designed to perform the task actively, with their intent set to $\theta_r = 0.001$.

4) Procedure: Upon arrival at the experimental site, participants first completed the consent form. Then, the staff demonstrated how to operate the joystick to push the object in the virtual scenario. Participants were divided into 14 pairs, and each pair was given a 5-min practice session before the experiment to become familiar with the collaborative process by collaborating with both human partners and robot agents.

During the experiment, two participants sat separately to prevent them from seeing each other's screens and motions. Each participant completed a total of 35 trials (5 trials for each HRC condition and 10 trials for the HHC condition). The trial conditions were randomly assigned by the staff and were not disclosed to participants. In the HRC conditions, each of the two participants collaborated with the robot agent on their respective PCs. In the HH condition, the screen of PC 2 was switched to display the content of PC 1. Human actions in PC 2 were transmitted via the internet cable to PC 1 and applied to the left end of the object. The screens of the two PCs displayed the same content to prevent participants at PC 2 from noticing the screen switch.

Participants were required to start the trial together and to complete the experiment within 1 min. Each participant was accompanied by a staff. To conclude a trial, participants assessed the object's position on the screen and released the joystick to stop the action. In the HRC conditions, completion times could vary significantly between participants. To prevent unintended partner identification, participants were required to remain seated quietly for the full one-minute duration, avoiding any sounds or movements that could indicate whether their partner was a human or a robot. When the time was up, the trial was concluded regardless of the completion status. After each trial, participants filled out Questionnaire 1 (shown in Table I) and exited the experimental area together until the researchers set up the experimental condition for the next trial.

Questionnaire 1 is used to collect participants' subjective evaluations of each experimental condition. The "partner-rating" item is used to measure the participant's satisfaction with the unknown collaborative partner. The "self-rating" item is used

to assess participants' satisfaction with their own operations. The "overall-rating" item is used to evaluate how much the goal-tracking accuracy of each trial meets the participant's expectations. Participants were required to rate the above items on a scale from 0 to 10, allowing up to one decimal place. The "classification" item is used to examine whether participants can accurately distinguish between human and robot collaborators.

5) Metrics: Quantitative metrics: Quantitative metrics compare the effects of HRC with that of HHC through the data recorded in the human subjects experiments and compare the performance of different robot control strategies.

Timestamps are recorded for each trial. The completion time of each trial is measured as follows:

$$\mathcal{T} = t_{\text{end}} - t_{\text{start}} \tag{22}$$

where t_{start} is the timestamp when one of the participants starts to take actions and t_{end} is the timestamp when both participants stop taking actions.

The metric measures the goal-tracking error for each trial using the following equation:

$$\mathcal{C} = \|x_{\text{end}} - x_{\text{goal}}\| \tag{23}$$

where $x_{\text{end}} \in \mathbb{R}^6$ is the state of the object when the trial is concluded. The smaller $\mathcal{C}$ is, the closer the object is to the target position.

The coordination of actions between two agents is measured with an alignment metric

$$\mathcal{H} = \frac{\sum_{k=0}^{\mathcal{L}} |\omega(k)|}{\mathcal{T}} \tag{24}$$

where $\mathcal{L}$ represents the length of each trial. The better the coordination between the agents is, the more aligned the ends of the object will remain during movement, leading to a smaller $\mathcal{H}$.

The number of times the human agents adjust their actions during the collaboration is counted according to the following equation in order to analyze the efforts of the human agents to adapt to their collaborating partners' actions in the different experimental conditions

$$\mathcal{N} = \sum_{k=0}^{\mathcal{L}-1} c(k)$$
$$c(k) = \begin{cases} 1 & \text{if} \quad |u_h(k+1) - u_h(k)| > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

In the HH condition, $\mathcal{N}$ is calculated as the mean of the number of times both human agents adjust their actions. The smaller $\mathcal{N}$ reveals that it is easier for the human agent to adapt to the partner's action.

In the experimental scenario, the collaborators need to work together to push the object to the goal. In each trial, the object moves at a low speed, and the desired displacement of the object remains constant. In addition, the action candidates set outlined in Section IV-B2 ensures no significant disparity in the magnitude of forces available to the collaborators. Therefore, it can be approximated that the two agents perform equal work to complete the task each time. The sum of the power of the two agents during the collaboration process is calculated to evaluate work efficiency

$$\mathcal{P} = \frac{\sum_{k=0}^{\mathcal{L}} |u_{\text{left}}(k) + u_{\text{right}}(k)|\, v_y(k) \cos(\omega(k))}{\mathcal{T}} \tag{26}$$

where $u_{\text{left}}(k)$ and $u_{\text{right}}(k)$ are the forces exerted by the two agents on the left and right ends of the object, respectively. The larger the $\mathcal{P}$ is, the more efficiently the task is accomplished.

In addition, a metric denoting the agent's role at the left end of the object is formulated as follows:

$$\alpha_{\text{left}}(k) = \begin{cases} 1 & \text{if} \quad |u_{\text{left}}(k)| - |u_{\text{right}}(k)| > 0 \\ 0 & \text{else if} \quad |u_{\text{left}}(k)| - |u_{\text{right}}(k)| < 0 \\ \alpha_{\text{left}}(k-1) & \text{otherwise.} \end{cases} \tag{27}$$

When $\alpha_{\text{left}}(k) = 1$, the left agent is the *leader*, whereas when $\alpha_{\text{left}}(k) = 0$, the left agent is the *follower*. The role of the right agent is denoted as α_{right}, which can be calculated in the same way as (27). Within the same trial, $\alpha_{\text{right}} + \alpha_{\text{left}} = 1$, indicating that when the left agent acts as the *leader*, the right agent assumes the *follower* role.

The proportion of time that the agent acts as the *leader* in the entire trial is represented as p_{lead}. According to the location of the agent, p_{lead} is divided into $p_{\text{lead}}^{\text{left}}$ and $p_{\text{lead}}^{\text{right}}$, which can be calculated as follows:

$$p_{\text{leader}}^{\text{left}} = \frac{\sum_{k=0}^{\mathcal{L}} \alpha_{\text{left}}(k)}{\mathcal{L}} \tag{28}$$

$$p_{\text{leader}}^{\text{right}} = 1 - p_{\text{leader}}^{\text{left}}. \tag{29}$$

p_{lead} denotes the general role of the agent. Inspired by information entropy [50], the frequency of role exchanges between the two agents throughout the collaborative process under different experimental conditions is analyzed as follows:

$$\mathcal{E} = -\left(p_{\text{leader}}^{\text{left}} \log p_{\text{leader}}^{\text{left}} + p_{\text{leader}}^{\text{right}} \log p_{\text{leader}}^{\text{right}}\right). \tag{30}$$

A smaller $\mathcal{E}$ indicates a lower frequency of role exchanges.

Subjective metrics: Subjective metrics analyze participants' satisfaction with their collaborating partners, their own performance, and task completion across experimental conditions by calculating the average ratings for items in Table I. Furthermore, participants' classifications of collaborating partners are tallied to determine, which action control policy is most likely to be perceived as human.

Synthesized metrics: The agent's classification of the collaborating partner results in four categories: "RCAH," "RCAR," "HCAH," and "HCAR." The synthesized metrics summarize the quantitative metrics according to these four categories to analyze the performance of the agent when perceived as a human and to identify the characteristics that might make the user believe that the collaborating partner is a human.

C. Experiment 2: Manipulating Soft Tissue

Indirectly positioning internal feature points on the surface of soft tissue to the desired locations by manipulating the edge of

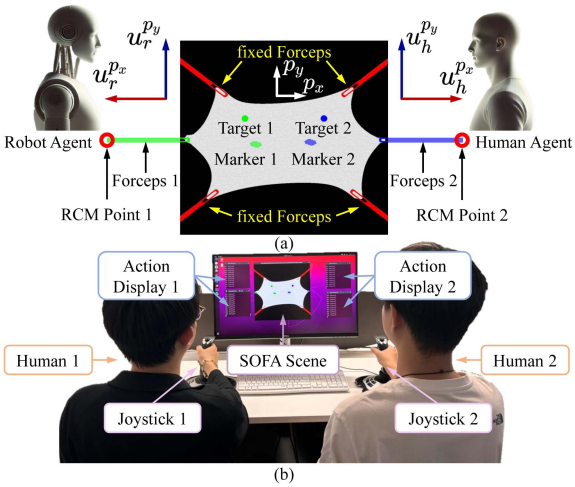

Fig. 4. Setup of Experiment 2. Human operators can stretch the soft tissue created by the SOFA with the virtual forceps through the joystick. The white coordinates are the reference coordinates for the scenario. The computer is used as an intelligent robot. The operator needs to stretch the tissue to the target position through HRC or HHC. (a) Scenario description for Experiment 2. (b) Experimental setup.

the tissue is a common task in laparoscopic surgery, enabling subsequent surgical tasks such as needle insertion or tissue dissection [51]. However, the highly nonlinear dynamic properties of soft tissues lead to unpredictable deformations, which complicates precise localization [24]. Experiment 2 validates the applicability of the proposed PNE strategy by extending it to this tissue manipulation task.

1) Hardware Setup: A virtual environment is constructed to simulate the surgical scenario using the simulation open framework architecture (SOFA), which is commonly employed to validate robot control algorithms in tissue manipulation tasks due to its high-quality simulated deformation behaviors of the soft materials [24], [52]. As shown in Fig. 4(a), soft tissue is modeled to represent human tissues in laparoscopic surgery, such as the kidney or peritoneum, and is fixed by four red forceps. The soft tissue can be stretched by the green and blue forceps, both of which are constrained by the remote center of motion points. The internal markers deform when the edge points are moved. The white coordinates serve as the reference coordinates for the scenario. Fig. 4(b) shows the experimental setup. In the HRC condition, the user manipulates a joystick to control one of the virtual forceps, and the robot, simulated by the computer, controls the other forceps to collaborate with the user. For HHC, two participants control the two forceps using separate joysticks. The inputs of the agents are the velocities of the forceps along the x- and y-axis. Participants need to control the joystick to pull the tissue at different velocities. The screen displays the current velocity of the forceps to guide users to adjust their actions. The joystick communicates with the SOFA through ROS.

2) Parameter Setting: The central positions of Marker 1 and Marker 2 are denoted as $p_1 = [p_{x_1}, p_{y_1}]^T$ and $p_2 = [p_{x_2}, p_{y_2}]^T$, respectively. Then, the state vector $x \in \mathbb{R}^4$ of the task can be defined as $x = [p_{x_1}, p_{y_1}, p_{x_2}, p_{y_2}]^T$. The initial state is set to $x_0 = [-15, 0, 15, 0]^T$. The central position of Target 1, $p_{t_1} = [-35, 20]^T$, and the central position of Target 2, $p_{t_2} = [35, 20]^T$, together form the goal state $x_{\text{goal}} = [-35, 20, 35, 20]$. The finite candidate set of agent actions representing the forceps control velocity is set to $U_i^j = \{-1.5, -0.9, -0.4, 0, 0.4, 0.9, 1.5\}$, where $j \in \{p_x, p_y\}$ represents the direction of the velocity. The penalty matrix Q_i is denoted as $\text{diag}(10, 10, 10, 10)$. The intent set describing the action preference of the agent is set to $\Theta_i^j = \{0.001, 1\}$. When $\theta_i^j = 0.001$, the agent prefers to manipulate the soft tissue quickly and track the designated target actively. When $\theta_i^j = 1$, the agent tends to manipulate the tissue slowly to avoid negatively impacting the overall tracking accuracy. Since Experiment 2 considers the agent is actions and intents in two directions, the cost function is formulated for each direction as follows:

$$C_i^j = \sum_{n=0}^{H-1} (\tilde{x}^T(k+n) Q_i \tilde{x}(k+n) + \theta_i^j (u_i^j(k+n))^T u_i^j(k+n)). \tag{31}$$

Because the robot agent needs to simultaneously infer the human agent's intents in two action directions, the joint intent space is expanded to $(\Theta_h^{p_x} \times \Theta_h^{p_y}) \times (\Theta_r^{p_x} \times \Theta_r^{p_y})$.

3) Experimental Tasks and Conditions: The primary purpose of this experiment is to evaluate the potential application of the proposed PNE action control policy in a complex HRC scenario. Specifically, the aim is to assess whether the PNE strategy can demonstrate good performance and ensure a good subjective evaluation. To achieve this, participants are required to move their respective markers as accurately as possible to the target points of the same color by manipulating the soft tissue.

Participants conduct the experiment under both HRC and HHC conditions. Under the HRC condition, the robot agent performs the task using both the proposed PNE policy and a baseline policy. The baseline is an ego nonempathic strategy, which is formulated as follows:

$$\min_{\xi_r^j \in \Xi_r^j} C_r^{\text{baseline}}(\xi_r^j) := C_r\left(\xi_r^j, \xi_{\text{obs}}^j, \theta_r^j\right) \tag{32}$$

where ξ_{obs}^j is the predicted human action trajectory based on the observed human action, i.e., $\xi_{\text{obs}}^j[0] = u_h(k-1)$. The intent of the baseline is set as $\theta_r^j = 0.001$ to enable the robot to actively perform the task during the collaboration. The robot using the baseline strategy plans its actions with a fixed intent, inferring human actions directly from observation with the assumption that the human knows the robot's intent. In other words, the baseline policy lacks empathy in collaboration.

4) Procedure: As in Experiment 1, participants familiarized themselves with the equipment under the guidance of the staff after signing the consent form. Participants were divided into 8 pairs and shown the relative positions of the markers and targets under a well-collaborated performance. Each pair was given five minutes to practice under both HRC and HHC conditions. Since the manipulated object was soft tissue, the actions of Agent 1 would cause the marker controlled by Agent 2 to move. Therefore, participants were reminded to avoid actions that could compromise their partner's target tracking accuracy as much as possible when adjusting their own actions.

TABLE II
QUESTIONNAIRE 2 OF THE HUMAN SUBJECTS EXPERIMENTS FOR SUBJECTIVE METRICS

	Question
Partner performance	Please assess your partner's performance.
Self-performance	Please assess your performance.
Overall performance	Please assess the goal-tracking error achieved in this collaboration.
Effort	Please assess the effort you have put in to achieve the performance you expect.
Frustration level	Please assess the frustration you felt during the collaboration.

Before the experiment, participants selected specific markers to manipulate and were required to continuously operate those markers until the experiment concluded. During the experiment, participants took turns performing HRC tasks. Each participant completed a total of 20 trials (5 trials in each HRC condition and 10 trials in the HHC condition). The trial conditions were assigned randomly by the staff. Participants were required to complete each trial within 1 min, with the option to decide whether or not to finish the trial based on the current positions of the markers and the goals. When the time reached 1 min, the trial ended regardless of completion status. After participants finished each trial, they were asked to fill out Questionnaire 2 shown in Table II, which is adapted from the NASA TLX questionnaire [53], and then leave the experimental area until the researchers set up the next experimental condition.

5) Metrics: Quantitative metrics: In the task of indirectly positioning soft tissues, target tracking accuracy is the primary metric of interest [24], [51]. The goal-tracking error $\mathcal{C}$ is defined using (23). In addition, the metric used to evaluate the goal-tracking error of agent i for its designated target is formulated using the following equation:

$$\mathcal{C}_{\text{agent}} = \|p_{z_{\text{end}}} - p_{t_z}\| \tag{33}$$

where $z \in \{1, 2\}$ represents the identifier of both the marker and the target. $p_{z_{\text{end}}} \in \mathbb{R}^2$ is the position of the marker z when the trial is finished.

The effort exerted by the agent i to achieve target tracking is defined as follows:

$$\mathcal{F}_i = \sum_{k=0}^{\mathcal{L}} \sqrt{((u_i^{p_x})^2 + (u_i^{p_y})^2)}. \tag{34}$$

A larger $\mathcal{F}_i$ indicates more effort that the agent puts into tracking the designated target. The sum of the effort of the human agent and the robot agent under HRC conditions is calculated as $\mathcal{F}_{\text{sum}} = \mathcal{F}_h + \mathcal{F}_r$.

Subjective metrics: Subjective metrics are collected through items in Table II. Participants rate their collaboration experience across five dimensions on a scale from 0 to 10, allowing up to one decimal place. The "partner performance" item is used to assess the degree of agreement between the robots' actions and participants' expectations. The "self-performance" item is designed to evaluate participants' satisfaction with their own operations. The "overall performance" item aims to evaluate how much the achieved goal-tracking accuracy meets participants' expectations. For the above three items, higher scores indicate better collaboration effectiveness. The "effort" item is used to evaluate participants' self-assessment of their effort while achieving the desired target tracking. A lower score indicates less effort exerted by the participant, suggesting that the collaboration is easier. The "frustration level" item is used to assess participants' negative emotions they felt during the task, with lower scores indicating fewer negative emotions.

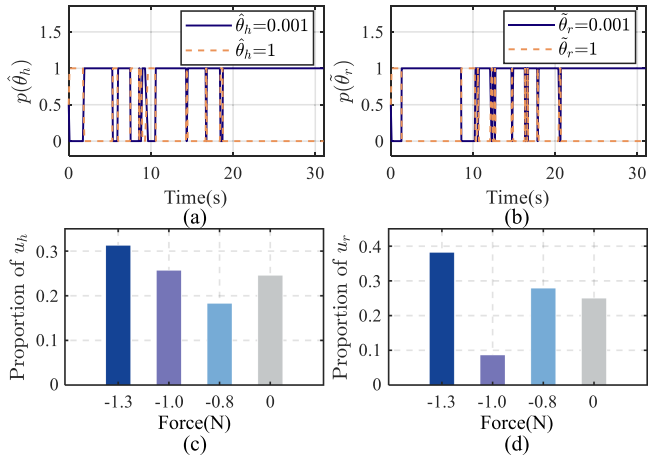

Fig. 5. Results of empathetic intent inference from the perspective of the robot and actions taken by the human and robot agents are shown in the figures. (a) Probability of the estimated human's intent under empathic inference. (b) Probability of the robot's inference regarding the human's estimated intent of the robot. (c) Proportions of each force over the human's action trajectory. (d) Force proportions over the robot's action trajectory.

V. RESULTS AND DISCUSSION

This section presents and analyzes the results of the two human subjects experiments, and discusses the findings observed. To enhance clarity, the previously defined abbreviations are listed in the Nomenclature.

A. Results of Experiment 1

1) Results of the Performance Test: Prior to the human subjects experiments, a performance test was conducted to evaluate the effectiveness and applicability of the empathetic inference-based method. During the test, the operator, positioned at the right end of the object, collaborates with the PNE robot agent. The goal was to maintain the object's alignment while moving and to ensure accurate positioning at the target, following the same procedure as in the human subjects experiments.

Fig. 5(a) and (b) presents the results of empathetic intent inference from the perspective of the robot. The proportions of each force adopted by the two agents from the set U_i over their entire action trajectories are computed, and the results of the calculations are shown in Fig. 5(c) and (d), respectively. According to Fig. 5(a) and (c), it is evident that the operator prefers to push the object with maximum force, from which the robot infers that the operator generally acts with an active intent. Meanwhile, Fig. 5(b) illustrates that in the robot's inference, the operator thinks that the robot's intent is generally active. As shown in Fig. 5(d), the robot adjusts its intent in alignment

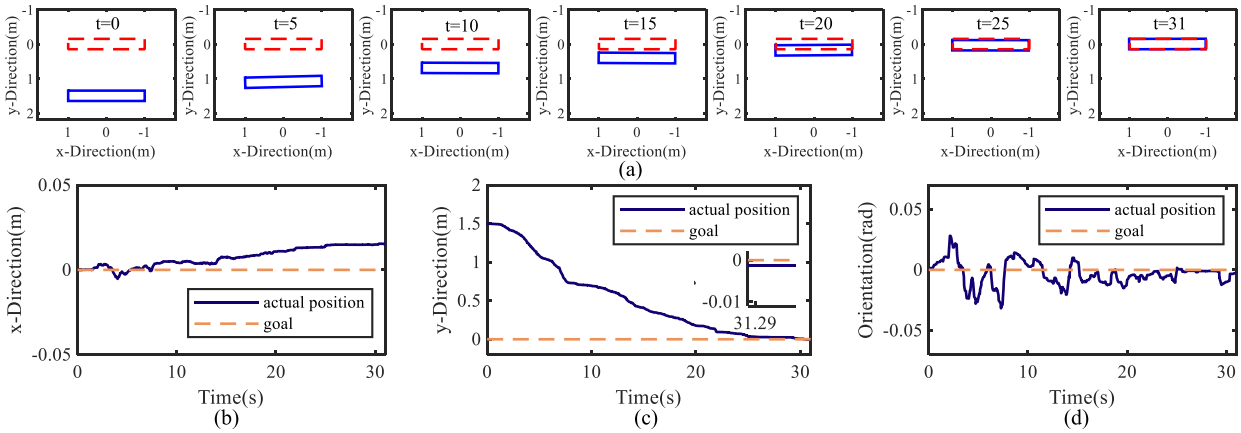

Fig. 6. Location snapshots and the trajectories of the object are plotted. (a) Snapshot of the object (solid blue rectangle) moving toward the goal (dashed red rectangle). (b) x-direction displacement of the object. (c) y-direction displacement of the object. (d) Orientation of the object.

with this inference and generally applies maximal force accordingly. The above results validate the applicability of the action preference defined in Section IV-B2, demonstrating that, in a practical environment, the set intent can effectively characterize the agent's action tendency.

Fig. 6 illustrates the snapshots and trajectories of the object motion during the collaboration process. The object remains well-aligned while moving forward and ultimately nearly coincides with the target position, suggesting that the proposed empathetic inference-based control policy enables the robot to achieve good collaboration with the operator.

2) General Results of Human Subjects Experiments: Quantitative results: The quantitative metrics of Experiment 1 under different experimental conditions, as described in Section IV-B5, are plotted as boxplots and presented in Fig. 7. The performance of different collaboration policies is compared based on the following dimensions.

With inference versus without inference: The human can adjust actions to adapt to different action strategies well to accomplish the task. As the metric $\mathcal{C}$ in Fig. 7(b) shows, the collaboration of the participant with the baseline policy is capable of achieving good goal-tracking accuracy. The t-test for $\mathcal{C}$ indicates that the baseline outperforms the RNE ($p < 0.05$) but does not show a significant difference compared to the other robot control strategies ($p > 0.1$). The boxplots for metric $\mathcal{N}$ indicate that participants frequently adjust their actions when collaborating with different robot agents. The results of the t-tests for $\mathcal{N}$ in Fig. 7(d) indicate no significant difference in the effort participants exert when adapting to the actions of robots with or without inference ($p > 0.05$). The above results suggest that regardless of whether the robot employs intent inference or follows a fixed-action strategy without inference, human participants can adjust their actions accordingly to ensure goal-tracking accuracy. Fig. 7(c) highlights that the baseline robot, which does not consider human actions, coordinates significantly worse with participants' actions compared to other designed robot agents ($p < 0.001$). In addition, since the human has to conform to the fixed manner of the baseline controller, the distribution of $\mathcal{T}$ in Fig. 7(a) is the most centralized among all of the HRC

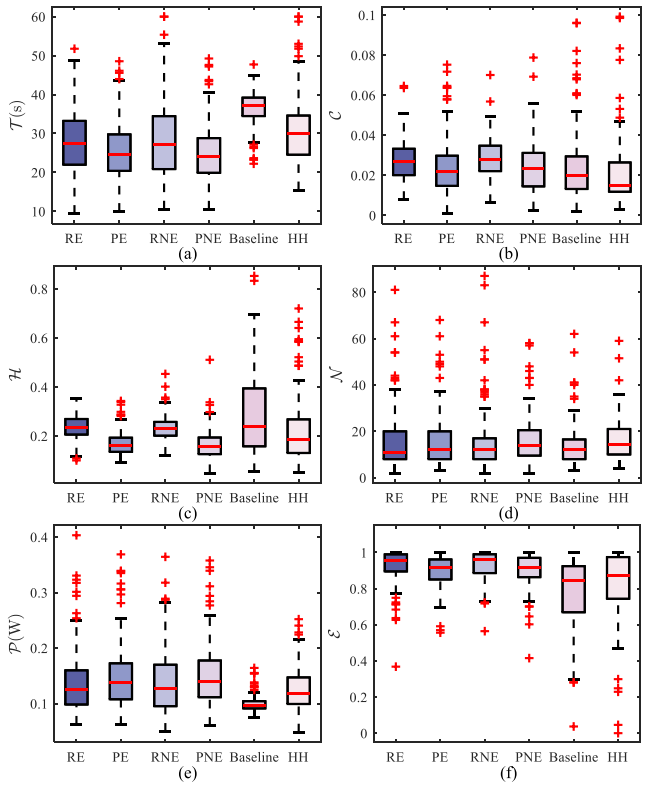

Fig. 7. Boxplots of quantitative metrics across experimental conditions. (a) Boxplots of completion times for each experimental condition. (b) Boxplots of goal-tracking error for each experimental condition. (c) Boxplots of alignment for each experimental condition. (d) Boxplots for the number of action adjustments under each experimental condition. (e) Boxplots of the summed power of the two agents under each experimental condition. (f) Boxplots of the role exchange frequency for each experimental condition.

conditions. This results in much lower collaborative efficiency for the baseline strategy than for other robot agents ($p < 0.001$), as illustrated in Fig. 7(e). Therefore, while human adaptability ensures goal-tracking accuracy, the fixed action approach of the baseline strategy imposes many limitations on achieving efficient and seamless collaboration.

TABLE III
RESULTS OF T-TESTS

	$\mathcal{C}$	$\mathcal{H}$	$\mathcal{N}$	$\mathcal{P}$	$\mathcal{E}$
PE vs. HH	p=0.060	p<0.001	p=0.499	p<0.001	p<0.001
PNE vs. HH	p=0.072	p<0.001	p=0.846	p<0.001	p<0.001
PE vs. PNE	p=0.853	p=0.756	p=0.645	p=0.875	p=0.544

Reactive versus proactive: It is critical for the robot agent to take into account the dependency of the human's planning on the robot's next action when planning its own actions. As shown in Fig. 7, proactive robot agents facilitate faster task completion compared to reactive robot agents during collaboration. Significant differences in $\mathcal{T}$ are observed in the t-tests for PE versus RE ($p < 0.05$), PE versus RNE ($p < 0.05$), PNE versus RE ($p < 0.01$), and PNE versus RNE ($p < 0.01$). As is clearly visible, the proactive policies can generally achieve smaller goal-tracking errors than the reactive policies. The t-tests for $\mathcal{C}$ reveal that tracking accuracy in proactive conditions is significantly higher than in reactive conditions (PE versus RE: $p < 0.05$; PE versus RNE: $p < 0.01$; PNE versus RE: $p < 0.05$; PNE versus RNE: $p < 0.001$). Furthermore, the t-tests for $\mathcal{H}$ show that proactive controllers significantly outperform the reactive controllers in terms of coordination with the human actions (PE versus RE: $p < 0.001$; PE versus RNE: $p < 0.001$; PNE versus RE: $p < 0.001$; PNE versus RNE: $p < 0.001$). In summary, proactive robot agents exhibit better adaptability to human actions than reactive agents, leading to more appropriate behavior and better overall performance.

HRC versus HHC: The frequency of role exchanges in HHC is more diverse than in the proposed HRC conditions. As the metric $\mathcal{E}$ in Fig. 7(f) shows, role exchange occurs in most trials across all experimental conditions. However, $\mathcal{E}$, which measures the frequency of role exchanges, is generally higher when humans collaborate with the proposed robotic action policies compared to HHC. The results of student t-tests conducted on the quantitative metrics in the PE, PNE, and HH conditions are shown in Table III. It is evident that $\mathcal{E}$ in HHC significantly differs from that in the PE and PNE conditions. In HHC, participants exhibit both low- and high-frequency role exchanges, leading to greater diversity in $\mathcal{E}$. In the baseline condition, $\mathcal{E}$ are significantly different from that in other HRC conditions ($p < 0.001$), which could indicate that humans could exchange roles with different agents through different frequencies.

The t-tests for metric $\mathcal{C}$ show that PE and PNE robots, when collaborating with humans, can achieve goal-tracking accuracy comparable to that of HHC. Furthermore, the p values of PE versus HH and PNE versus HH in metrics $\mathcal{H}$ and $\mathcal{P}$, along with Fig. 7(c) and (e), suggest that PE and PNE robot agents significantly improve the coordination with the human actions, and the collaborative efficiency, compared to the HHC.

Ego versus nonego: A nonego robot agent can further improve the collaborative effectiveness compared to an ego robot agent. The t-test results for the metrics in Table III reveal no significant differences between the PE and PNE control policies, implying that, from a statistical perspective, the ego robot and the nonego robot have similar capabilities. However, Fig. 7(b), (c), and (e) shows that the PNE robot agent generally performs better than the PE robot agent. The PNE policy can coordinate with humans better and achieve more accurate and efficient performance than the PE policy.

TABLE IV
MEANS AND STD OF SUBJECTIVE RATING ITEMS UNDER DIFFERENT CONDITIONS

Condition	Partner-Rating	Self-Rating	Overall-Rating
RE	8.6921± 0.9338	9.0650± 0.9253	8.8036± 0.8875
PE	8.8721± 0.9165	8.9900± 0.8747	8.9407± 0.9043
RNE	8.5614± 1.0733	8.9979± 0.8458	8.6550± 0.8091
PNE	8.9786± 0.9154	9.1150± 0.8469	9.0450± 0.9164
Baseline	8.5321± 1.2519	9.0650± 0.8976	8.9264± 0.9026
HH	9.2400± 0.8163	9.2904± 0.7645	9.3675± 0.6717

TABLE V
PARTICIPANTS' IDENTIFICATION OF COLLABORATING PARTNERS

Condition	RCAR	RCAH	Percentage of RCAH
RE	91	49	35.00%
PE	75	65	46.42%
RNE	102	38	27.14%
PNE	75	65	46.42%
Baseline	95	45	32.14%
Condition	**HCAR**	**HCAH**	**Percentage of HCAH**
HH	88	192	68.57%

Subjective results: The mean values and standard deviations (STD) of the subjective rating items from Table I are presented in Table IV. The partner-rating shows that the human is the most preferred collaborating partner. Participants express the highest confidence in their own performance and the greatest satisfaction with task completion when working with human partners, as reflected in the self-rating and overall-rating. Among the proposed control policies, the PNE control strategy has the highest ratings in all items, which implies that a proactive and nonego robot agent is the most favored robot for participants.

The PE and PNE robot agents exhibit human-like behaviors during collaboration. Table V presents the number of times the collaborating partner is classified as a human or a robot in each experimental condition. In the HHC condition, participants perform a total of 280 recognitions. It can be seen that the probability of a human being classified as human is higher than the probability of a robot being classified as human. Notably, under the PE and PNE conditions, the robot agent is significantly more frequently classified as a human than under other HRC conditions. This indicates that not all robot agents can be easily thought to behave like humans in the designed experimental scenario, which suggests that the classification results make sense. Therefore, the classification results empirically show that the PE and PNE robot agents' actions can make participants feel as though they are collaborating with human partners. This suggests that when a robot considers the optimal human actions corresponding to its future actions, it is perceived as more human-like from the participant's perspective.

Summary: The general analysis of quantitative and subjective metrics indicates that the PNE robot agent demonstrates strong performance when collaborating with different participants. It closely approximates participants' perceptions of a human partner and achieves the highest collaborative effectiveness and user satisfaction among the HRC conditions. This suggests that

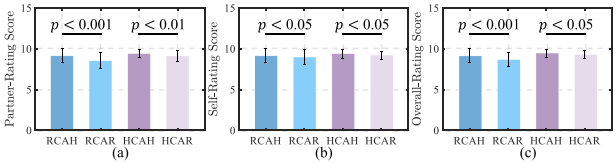

Fig. 8. Means of subjective scores corresponding to the classification categories in HRC and HHC conditions. These results reveal that participants think that human partners could perform better than robot partners. (a) Results of partner-rating. (b) Results of self-rating. (c) Results of overall-rating.

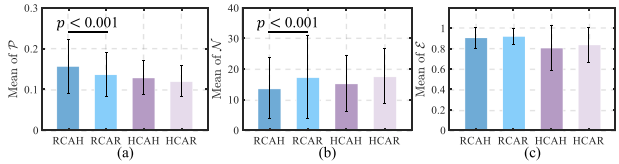

Fig. 9. Under RCAH, RCAR, HCAH, and HCAR categories, the mean values of the summed power of both agents, the number of action adjustments, and the frequency of role exchanges in the HRC and HHC conditions. (a) Average values of the summed power. (b) Average number of action adjustments. (c) Average values of role exchange frequency.

the PNE action control policy enables the robot to behave like a human. Moreover, while HHC achieves good goal-tracking accuracy, the PNE robot agent outperforms it in terms of collaboration efficiency and coordination. This highlights the potential of the PNE control policy in scenarios where fluid collaboration is crucial, such as jointly transporting an object on a large tray. Based on the synthesis metrics, the following context will study the characteristics that could make the participant believe the partner is a human.

3) Hypothesis Testing: To analyze the association between classification results and user evaluations, the means and STDs of participants' subjective ratings are calculated for instances where collaborating partners are categorized into RCAH, RCAR, HCAH, and HCAR entries under HRC and HHC conditions. To ensure consistency in comparisons, the 62 trials in which both agents correctly classified their partners as human out of 140 trials in the HH condition are categorized as HCAH. The baseline policy is excluded from the HRC condition. As illustrated in Fig. 8, participants consistently preferred partners identified as humans over those identified as robots in both HRC and HHC conditions. The partner-rating results show that human agents are perceived to perform better than robot agents. The self-rating results reveal that participants are more satisfied with their performance when they perceive their partners as humans rather than robots. The overall-rating results indicate that participants believe human partners achieve better goal-tracking accuracy than robot partners. These findings empirically indicate that human-like behavior, under subjective cognition, enhances user evaluations. Therefore, identifying the behavioral characteristics that lead users to perceive their collaborating partner as human is crucial for developing human-like behavioral strategies in HRC. To address this, this section also tests a series of hypotheses using synthesized metrics.

The metrics (25), (26), and (30) formulated in Section IV-B5 are implicit collaborative performances that participants cannot directly observe. The following hypotheses are tested to examine whether these metrics affect the classification results.

Hypothesis 1 (H1): When there is a time limitation, the agent that accomplishes collaborative tasks with the user efficiently is more likely to be perceived as human.

Hypothesis 2 (H2): When the user exerts less effort in adjusting actions during the collaboration process, the collaborating partner is more likely to be perceived as human.

Hypothesis 3 (H3): When role exchanges occur less frequently throughout the collaboration, the user is more likely to perceive the partner as human.

Results: To test the above hypotheses, the means of the summed power $\mathcal{P}$, the action adjusting number $\mathcal{N}$, and the frequency of role exchange $\mathcal{E}$ under RCAH, RCAR, HCAH, and HCAR in HRC and HHC conditions are plotted in Fig. 9. As illustrated in Fig. 9(a), the summed power is higher when participants classify their collaborating partners as human rather than robot. This reveals that participants believe that a human agent can accomplish the collaboration more effectively than a robot agent, especially when there is a one-minute time limit (supporting *H1*). The means of metric $\mathcal{N}$ under the RCAH, RCAR, HCAH, and HCAR categories presented in Fig. 9(b) suggest that, regardless of the collaborative policy employed, participants spend less effort in adjusting actions to match their partners' actions when they classify their partners as humans (supporting *H2*). It can be empirically analyzed that participants believe the human agent can take more understandable as well as more predictable actions than the robot agent. As a result, participants think they can collaborate more easily with human partners. As for *H3*, Fig. 9(c) shows that although no significant difference is observed, the means of $\mathcal{E}$ is overall lower when the partner is identified as a human than when the partner is identified as a robot (partly supporting *H3*). This implies that in the lack of preassigned roles, the role exchange frequency throughout the collaboration might have some impact on the outcome of the subjective judgment, but it is not the primary feature for participants when classifying partners.

The goal-tracking accuracy and the alignment of the object, as defined by metrics (23) and (24), are explicit collaborative performances that participants can directly observe. The goal-tracking accuracy depends on three parameters: the final position of the object in the x-direction, the final position of the object in the y-direction, and the object's final orientation. Two additional hypotheses are proposed to examine which location parameter matters for participants' classification results and the influence of the coordination of actions on classification, respectively.

Hypothesis 4 (H4): The user mainly focuses on the final position of the object in the y-direction. High tracking accuracy in the y-direction makes the user perceive the collaborating partner as human.

Hypothesis 5 (H5): When the two ends of the object maintain good alignment during the collaboration, the user is more likely to perceive the collaborating partner as human.

Results: To address *H4* and *H5*, the boxplots of the final state of the object and the alignment in RCAH and RCAR categories under the PNE condition, as well as in HCAH and

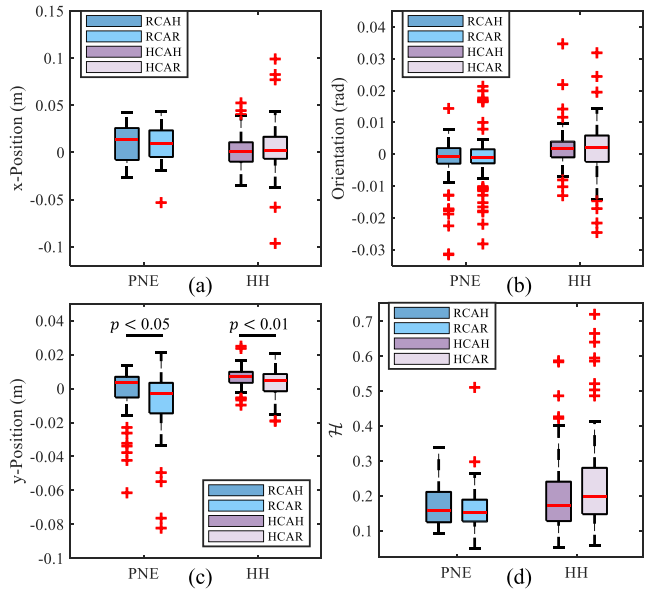

Fig. 10. Boxplots of the final state of the object and the alignment in RCAH and RCAR categories under the PNE condition, as well as in HCAH and HCAR categories under the HH condition. (a) Boxplots of the final position in the x-direction over PNE and HH experimental conditions. (b) Boxplots of the final position in the orientation for PNE and HH experimental conditions. (c) Boxplots of the final position in the y-direction under PNE and HH experimental conditions. (d) Boxplots of alignment in PNE and HH conditions.

HCAR categories under the HH condition, are shown in Fig. 10. Fig. 10(a) and (b) demonstrates that in both PNE and HH conditions, the object's final position in the x-direction and the object's final orientation when participants classify their partners as humans are not significantly different from those when participants classify their partners as robots. Fig. 10(c) indicates that the goal-tracking accuracy in the y-direction of the object when participants classify their partners as humans is significantly higher than that when participants classify their partners as robots under both experimental conditions. These results indicate that the user is more concerned about the tracking accuracy in the y-direction than in the x-direction and the orientation when classifying the partner. When the object is very close to the target position in the y-direction, the user perceives the collaborating partner as a human (supporting H4). Regarding H5, the boxplots in Fig. 10(d) illustrate that, regardless of the experimental condition, alignment when participants classify their partners as humans is not significantly different from the alignment when participants classify their partners as robots. In the PNE condition, alignment is generally better when the robot is classified as a robot than when the robot is classified as a human. This suggests that alignment is not a key feature for users to judge the type of collaborating partners (rejecting H5).

In sum, hypothesis testing for H4 and H5 demonstrates that the object's final position in the y-direction is an important characteristic for the participant to identify the partner. To further explore the target tracking feature that aligns with users' cognition of human behavior, a comparison of the object's final state between the RNE and PNE conditions is conducted. As Fig. 11(a) shows, the final position of the object in the x-direction

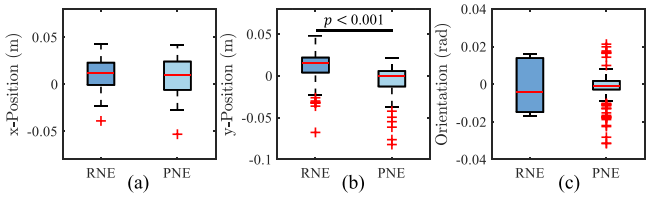

Fig. 11. The boxplots of the final state of the object in RNE and PNE conditions. (a) Boxplots of the final position in the x-direction. (b) Boxplots of the final position in the y-direction. (c) Boxplots of the final orientation.

in the RNE condition with the lowest RCAH percentage does not differ statistically from that in the x-direction in the PNE condition with the highest RCAH percentage. For the object's final orientation shown in Fig. 11(c), there is also no significant difference between the RNE and PNE conditions. However, Fig. 11(b) illustrates that the object's final position in the y-direction in the RNE condition is significantly different from that in the PNE condition. It can be seen that in the PNE condition, the object can be moved above the target position in the y-direction, whereas in the RNE condition, the object overall stops below the target position. These results indicate that in users' subjective cognition, the human's collaborative behavior can allow the object to exceed the target position in the primarily focused y-direction, but cannot stop the object before it reaches the goal position. This explains why a robot agent adopting the PNE policy is more likely to be recognized as a human than a robot agent using the RNE policy.

In sum, hypothesis testing results demonstrate that participants could believe that, compared to the robot agent, the human agent should perform more efficiently, and take actions that are easier to adopt. Furthermore, in the subjective perception of participants, human collaborative behavior could achieve greater tracking accuracy in the y-direction of the target position, and could demonstrate a more active pattern, allowing the object to slightly exceed the target position, compared to the collaborative behavior of the robot.

The conclusions from hypothesis testing and the general results shown in Fig. 7 together suggest that the proposed PNE robot achieves efficient, accurate, and active collaboration. This aligns with human behavior as perceived by participants, leading to the highest user satisfaction and the greatest confusion in distinguishing the robot from a human partner. Therefore, the PNE policy is further shown to enable the robot to exhibit human-like action characteristics. The classification results in Table V show that humans have a higher probability of being recognized as humans during collaboration, suggesting that human partners may exhibit additional distinguishing characteristics. Since Fig. 7(f) shows that the human exhibits the diversity in metric $\mathcal{E}$ when collaborating with different partners, the following context will further investigate this characteristic of the HH team.

The average proportion of time the agent acts as the *leader* during the collaboration for each experimental condition is presented in Fig. 12. Agents 1 and 2 refer to the right and left participants in the HH condition and to the human and robot in other conditions. This representation method will be used consistently in the following section.

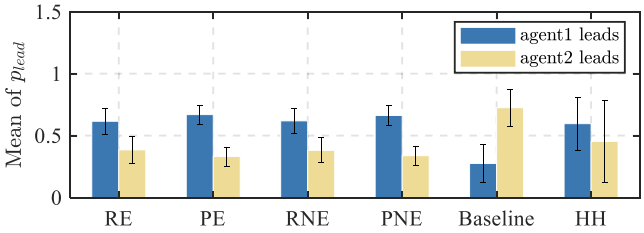

Fig. 12. Average proportion of time the agent acts as the *leader* for each experimental condition. Agents 1 and 2 denote the two participants under the HH condition, and the human and the robot, respectively, under other conditions.

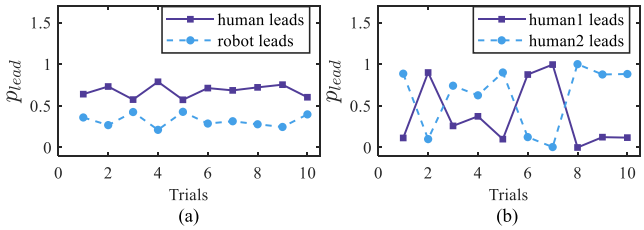

Fig. 13. Proportion of time the agent acts as the *leader* in HRC and HHC conditions for pair 14. (a) Proportion of time the agent acts as the *leader* when two participants collaborate with the PNE robot. (b) Proportion of time the participant acts as the *leader* during HHC.

Fig. 12 shows that in the HRC conditions employing the proposed policies, the participant primarily assumes the *leader* role. When the robot adopts the baseline policy, it can take actions with maximum force at the beginning of the task more quickly than the human and does not reduce thrust until the object is close to the goal position, without considering alignment. This policy makes the human act mainly as a *follower* when collaborating with the baseline robot. The above results indicate that humans can engage in collaboration as either *leader* or *follower*. The 14th pair of participants is chosen to illustrate the p_{lead} during collaboration in PNE and HH conditions. The general role in Fig. 13(a) shows that in HRC, the human is the *leader* from the whole in each trial. From Fig. 13(b), it can be seen that for the same pair of participants, Human Agent 1 can act as a *leader* in this trial but assume a *follower* in the next trial, which suggests that the same human agent can play different roles in the designed scenario. This characteristic could lead to a variety of $\mathcal{E}$ and average the proportion of time the agent acts as the *leader* from the whole, as everyone is able to act as a *leader* or a *follower* during the collaboration.

4) Case Study: A case study is presented to demonstrate the characteristics of role exchanging in the trials of HRC and HHC. The case chosen from HRC is the trial with the highest tracking accuracy under the PNE condition when the human agent classifies the robot agent as human. In this trial, the human agent exerts a force on the left end of the object. The trial that is selected under HHC is the one with the highest tracking accuracy when both participants recognize their partners as humans. The purpose of selecting these two cases is to investigate how the frequency of role exchange between agents differs under these two conditions when target tracking accuracy is high. The performance of the agents in the two cases is shown in Fig. 14. It is evident that

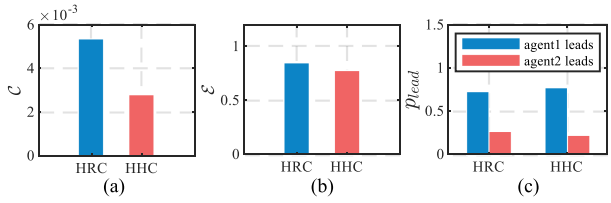

Fig. 14. General performances of the two agents in HRC and HHC cases. (a) Tracking error. (b) Frequency of role exchanges throughout the collaboration. (c) Proportion of time the agent acts as the *leader* during the collaboration.

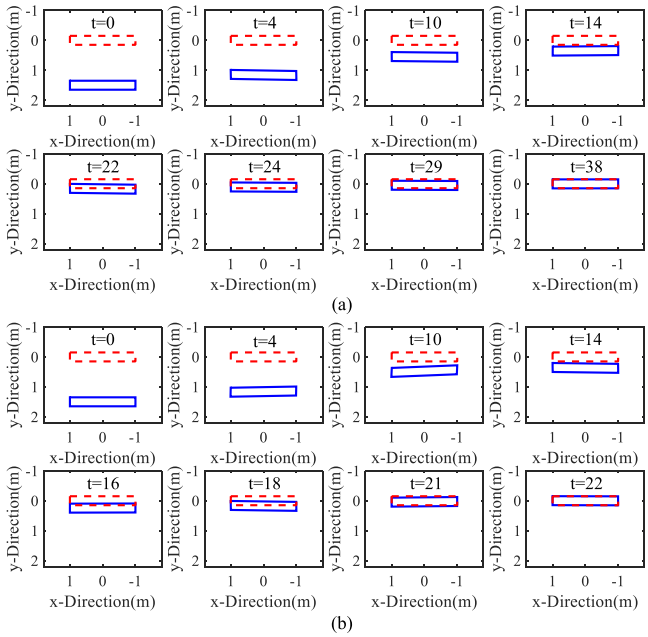

Fig. 15. Snapshots of the object (solid blue rectangle) moving toward the goal (dashed red rectangle) for the case study. (a) Object motion under HRC. (b) Object motion under HHC.

the chosen HH team achieves better tracking accuracy than the chosen human–robot team. Fig. 14(b) and (c) shows that the frequency of role exchange throughout the collaboration between the two teams has no noticeable difference.

Fig. 15(a) illustrates the object motion under the collaboration of the human agent and the PNE robot agent. As observed, the human primarily acts as the *leader* throughout the trial, regardless of the object's movement. Fig. 15(b) depicts object motion in HHC. At the beginning, the participant at the right end leads the task. However, as the object approaches the target position, the two agents gradually move it toward the goal by continuously exchanging roles.

The trajectories for both cases are plotted in Fig. 16 for a detailed comparison. The time at which collaboration ends in both cases is given in Fig. 16(a). Fig. 16(b) shows that, compared to the human–robot team, the HH team demonstrates a more active performance, bringing the object closer to the goal position in the mainly concerned y-direction. Fig. 16(c) illustrates that the coordination of actions between agents under the PNE condition ($\mathcal{H} = 0.1297$) is better than in HHC ($\mathcal{H} = 0.1842$).

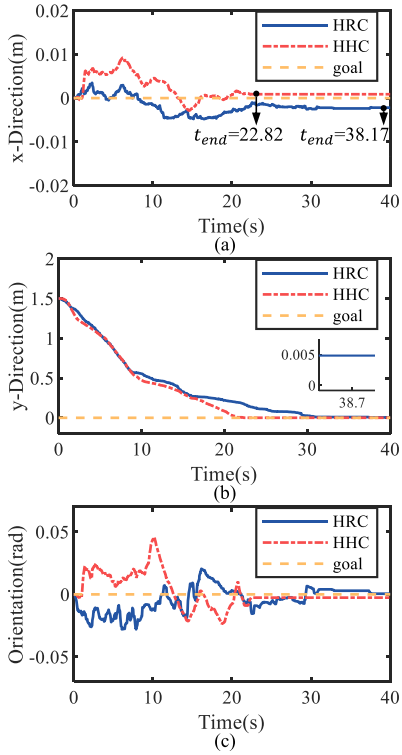

Fig. 16. Comparisons of object trajectories under case HRC and case HHC are plotted. (a) x-direction displacement of the object. (b) y-direction displacement of the object. (c) Orientation of the object.

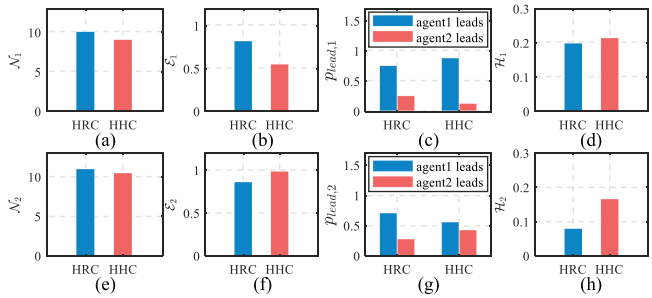

Fig. 17. Number of action adjustments, role exchange, and alignment in different phases for case HRC and case HHC are plotted. (a) Number of action adjustments in the rapid operation phase. (b) Frequency of role exchanges during the rapid operation phase. (c) Proportion of time the agent acts as the *leader* in the rapid phase. (d) Coordination of actions between agents in the rapid phase. (e) Number of action adjustments in the delicate operation phase. (f) Frequency of role exchanges in the delicate operation phase. (g) Proportion of time the agent acts as the *leader* in the delicate phase. (h) Coordination of actions between agents in the delicate phase.

As suggested in [54], the joint object transporting task can generally be divided into two phases: a rapid operation phase, during which the two agents move the object quickly toward the target when it is far from the goal, and a delicate operation phase, during which the agents slow down to precisely position the object as it approaches the target. This transition indicates that human attention to target-tracking accuracy intensifies as the object approaches the goal. We observed that in both cases, the average speed of the object after $p_y = 0.3$ is significantly lower compared to the average speed before $p_y = 0.3$. Building on the insights from this study, the collaboration process before $p_y = 0.3$ is defined as the rapid operation phase, while the collaboration process after $p_y = 0.3$ is defined as the delicate phase. This distinction allows for a detailed investigation of the dynamics of role exchanges during each phase. Metrics $\mathcal{N}_1$, $\mathcal{E}_1$, $p_{\text{lead},1}$, and $\mathcal{H}_1$ are introduced to quantify the number of action adjustments, the frequency of role exchanges, the proportion of time the agent acts as the *leader*, and the coordination of actions between agents, during the rapid operation phase. Similarly, metrics $\mathcal{N}_2$, $\mathcal{E}_2$, $p_{\text{lead},2}$, and $\mathcal{H}_2$ are introduced to quantify the same variables during the delicate operation phase.

Fig. 17(a) and (e) compares the effort exerted by the human agent adapting to the partner's action for both cases in different phases. As is visible, the human agent exerts more effort in adapting to the partner's actions in the HRC case than in the HHC case. For the same case, there is no noticeable variation in the number of action adjustments of the human agent in the two phases.

The comparison of Fig. 17(b) and (f) shows that when a human agent collaborates with the robot agent, there is no noticeable difference between $\mathcal{E}_1$ and $\mathcal{E}_2$. The role exchange pattern is consistent in that the human acts as the *leader* most of the time. But when two human agents collaborate, there is an obvious increase in role exchange frequency in the delicate operation phase. Combining Fig. 17(c) and (g), it is evident that when performing HHC, the roles of the two agents are explicitly divided into *leader* and *follower* in the rapid operation phase, resulting in a small $\mathcal{E}_1$. In contrast, role exchange becomes more frequent in the delicate operation phase, leading $\mathcal{E}_2$ to be much higher than $\mathcal{E}_1$. The analysis of Fig. 17(b), (f), (c), and (g) indicates that the HH team can make the frequency of role exchanges in the delicate operation phase significantly higher than that in the rapid operation phase. Fig. 17(d) and (h) shows that although the HH team and the human–robot team perform similarly in terms of metric $\mathcal{H}_1$ during the rapid operation phase, in the delicate operation phase, the human–robot team can achieve much smaller $\mathcal{H}_2$ than the HH team. This result suggests that although the frequency of role exchanges in HH teams can be significantly increased in the delicate operation phase, it may not necessarily guarantee coordination with the actions of the partner. Compared to the human, the robot can respond more appropriately to the actions of its partner, enabling more coordinated collaboration.

In sum, the case study demonstrates that achieving high goal-tracking accuracy, whether in HRC or HHC, requires the human agent to exert sufficient effort in matching the partner's actions. The HH team expects the object to move quickly in the rapid phase and then move slowly in the delicate operation phase. The designed robot agent can also exhibit this characteristic in collaboration with the human. Although Fig. 14(c) shows similar role exchange frequencies in HRC and HHC throughout the entire collaboration, the selected human–robot and HH teams demonstrate distinct role exchange patterns in different phases. In HRC, the role exchange frequency remains relatively consistent, regardless of the phase of operation. In HHC, as human attention to target-tracking accuracy intensifies

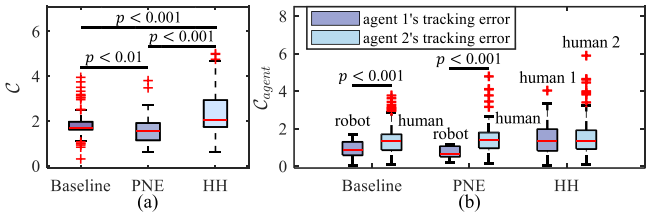

Fig. 18. Goal-tracking errors in each experimental condition are shown in boxplots. (a) Boxplots of goal-tracking error for each experimental condition. (b) Boxplots of goal-tracking errors achieved by the agents while tracking their respective targets under different conditions. Agents 1 and 2 represent the two participants under the HH condition, and represent the human and the robot under other conditions.

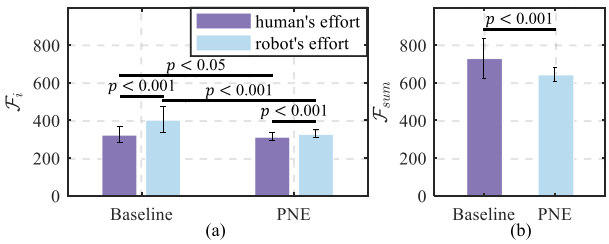

Fig. 19. Average values of the efforts exerted by agents during collaboration. (a) Mean efforts exerted by the human and the robot under the two HRC conditions. (b) Average of the summed efforts exerted by the human and the robot under the two HRC conditions.

when the object nears the target, the two agents significantly increase the frequency of role exchanges during the delicate operation phase. As a result, the selected HH team can achieve more flexible collaboration with a more active performance in the mainly concerned y-direction than the selected human–robot team. This may explain why the human is more likely to be classified as human than the robot. However, the increase in the frequency of role exchanges may reduce coordination with the partner's actions in the delicate operation phase.

B. Results of Experiment 2

Quantitative results: The boxplots illustrating the quantitative metrics $\mathcal{C}$ and $\mathcal{C}_{\text{agent}}$ from Experiment 2 across various experimental conditions are shown in Fig. 18. The mean efforts of the human agent and the robot agent, as well as the mean of the summed effort, are depicted in the bar chart shown in Fig. 19. Student's t-tests are performed on these quantitative metrics, and the results with significant differences are presented in the figures. The application potential of the proposed policy is validated through the following quantitative comparisons and subjective results.

Robot versus human performance: The robot agent outperforms the human agent in achieving good goal-tracking accuracy when manipulating the soft tissue. Fig. 18(a) indicates that the goal-tracking error achieved by the human–robot team is significantly smaller than that of the HH team. Fig. 18(b) shows the goal-tracking error achieved by the agent when tracking the designated target under different conditions. Agents 1 and 2 represent the two participants under the HH condition and represent the human and the robot under other conditions.

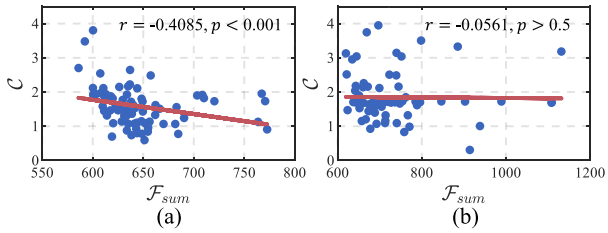

Fig. 20. Spearman correlation between goal-tracking errors and summed efforts under two HRC conditions. (a) Goal-tracking errors compared with the summed efforts under the PNE condition. (b) Goal-tracking errors compared with the summed efforts under the baseline condition.

It can be observed that both the baseline and the PNE robot agents achieve higher tracking accuracy compared to human agents when tracking their respective targets. Consequently, the overall target tracking performance of the human–robot team is significantly better than that of the HH team.

PNE versus baseline: Compared to the nonempathic baseline robot, the empathic PNE robot is a better collaborative partner, enabling the human agent to exert less effort to track the target. As shown in Fig. 19(a), regardless of whether the baseline strategy or the PNE strategy is adopted, the robot exerts more effort than the human to track the designated target. Combined with the human performance in Fig. 18(b), it suggests that sufficient effort is required to achieve good tracking accuracy.

Fig. 19(a) also indicates that both the human and the robot exert significantly more effort under the baseline condition than under the PNE condition. This leads to significantly higher summed effort under the baseline condition than under the PNE condition, as shown in Fig. 19(b). Fig. 18(a), however, demonstrates that human–robot teams in the PNE condition achieve smaller goal-tracking errors compared to those in the baseline condition. By examining the Spearman correlation between the summed efforts and the goal-tracking errors for both HRC conditions, there is a significant negative correlation between them ($r = -0.4085$, $p < 0.001$) under the PNE condition, as shown in Fig. 20(a). In contrast, Fig. 20(b) shows no significant correlation in the baseline condition ($r = -0.0561$, $p > 0.5$), indicating that agents exert redundant effort that does not positively contribute to tracking accuracy. It can be concluded from the quantitative results that the baseline policy causes agents to exert excessive effort, which reduces overall target tracking accuracy. In contrast, the PNE policy makes collaboration more effortless while ensuring good overall tracking accuracy.

Subjective results: Fig. 21 presents the mean and STD of participants' subjective ratings for the items in Experiment 2. It is visible that participants perceive the PNE robot partner as achieving the best performance and enabling the most effortless and least frustrating collaboration among all experimental conditions. According to the results of Student's t-tests in Fig. 21(a), participants perceive the actions of the PNE robot as significantly more aligned with their expectations than the baseline robot. As illustrated in Fig. 21(d) and (e), the PNE robot reduces users' perceived effort and frustration compared to the baseline robot.

Analysis: The quantitative results in Fig. 19 and the subjective results in Fig. 21 together demonstrate that although the

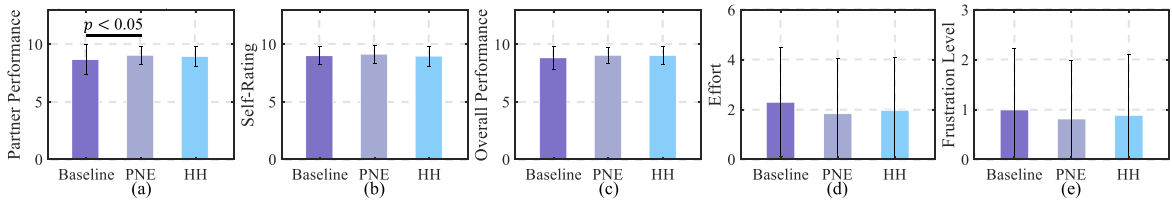

Fig. 21. Subjective results in each experimental condition are shown. (a) Average score of the partner's performance under each condition. Among all experimental conditions, the PNE robot partner receives the highest subjective ratings. (b) Average self-assessment scores of participants' performance under different conditions. (c) Average scores of the overall tracking errors under different conditions. (d) Average scores of participants' self-assessment of effort exerted in the collaboration under different conditions. (e) Average scores of participants' frustration levels under each condition.

ego baseline robot puts considerable effort into achieving good tracking accuracy with an active action pattern, its actions may be overly aggressive, potentially negatively affecting the human agent's desired tracking performance on the designated target due to the lack of empathy. For example, when both markers reach the vicinity of the target, the human agent may expect to pull the tissue slowly to complete the tracking, but the baseline robot may continue to pull rapidly. This action of the baseline robot could indirectly lead to an unexpected displacement of the human-operated marker, forcing the human to spend extra effort to readjust the marker to the desired location. In such cases, the human agent may, in turn, disrupt the robot's original tracking effect, consequently leading both agents to compensate for each other's errors by exerting excessive effort. It can be concluded that a baseline robot without empathy might prioritize target tracking accuracy over aligning its actions with human expectations, causing a mismatch between its behavior and human intent. This mismatch could cause the human and the baseline robot agents to interfere with each other and exert excessive effort, which ultimately reduces overall target tracking accuracy and user evaluation. In contrast, the robot adopting the PNE policy can adapt its actions to better align with human expectations and operational preferences by leveraging the results of empathetic inference, $\hat{\theta}_h^j$ and $\tilde{\theta}_r^j$. Therefore, although the PNE robot takes more actions than the human agent, it does not excessively interfere with the human's operation. Instead, it minimizes the human's redundant effort while maintaining good tracking accuracy and user evaluation.

Summary: From Fig. 18(a) and (b), it can be analyzed that even if good goal-tracking performance is shown to the participant, the effectiveness of HHC in a complex task is not always ideal. Integrating the robot's precise goal-tracking capability with the human's decision-making ability can lead to more accurate operations than those performed manually in a complicated collaborative task. However, the objective of a collaborative robot is not only to enhance the accuracy of operations but also to reduce the workload of users [24]. Figs. 19 and 20 illustrate that the empathic PNE strategy enables the operator to exert less effort compared to the ego nonempathic strategy. Fig. 21 indicates that, compared to the baseline policy, users perceive the PNE policy as offering a better collaboration experience. Combining quantitative and subjective results, it can be empirically demonstrated that the PNE policy enables the robot to be a good collaborative partner, improving task accuracy and user evaluation by exhibiting desired behavior. Taken together, we show that the proposed PNE policy has the potential to be generalized to complex tasks that require intent inference. In a practical scenario, the robot can adopt the proposed policy to operate the forceps at varying velocities, enabling collaboration with the surgeon to accomplish soft tissue manipulation tasks.

C. Discussion

In Experiment 1, the classification results presented in Table V indicate that, compared to reactive and baseline policies, the performance of the PNE robot aligns more closely with the user's subjective understanding of human behavioral characteristics. Although the HH team can make the frequency of role exchanges in the delicate operation phase much higher than that in the rapid operation phase, which might make agents more likely to be recognized as humans, this characteristic may not necessarily contribute to achieving good coordination with partners' actions. Since the primary objective of a collaborative robot is to fulfill seamless and effective collaboration, it does not need to consider making the robot exchange roles with the collaborating partner in a similar way that humans do. The metric $\mathcal{H}$ in Experiment 1 confirms that the coordination with human actions in the PNE condition is significantly better than in the HH condition. Experiment 1 demonstrates that human agents can achieve accurate target tracking in a simple task. However, Experiment 2 shows that as task complexity increases, human performance declines, and the involvement of robots improves task completion accuracy. Moreover, while the nonempathic ego robot in Experiment 2 achieves higher task accuracy than HHC by inferring human actions directly from observation, its fixed intent may lead the robot actions and the human actions to negatively affect each other, ultimately causing a decline in collaborative effectiveness and user evaluation in a complex task. By integrating findings from both experiments, it is evident that the PNE policy enables the robot agent to achieve a better performance than the human and the other adopted policies in diverse collaborative tasks. Therefore, it can be concluded that for effective collaboration, it is necessary for the robot to infer the intent of the human emphatically and plan actions by considering the dependence of the human's plan on the robot's next action. It also needs to dynamically adjust its intent based on empathetic inference throughout the collaboration. These capabilities enable the robot to exhibit human-like behavior, ensuring well-performed collaboration while enhancing user evaluation.

We acknowledge a limitation of the proposed method. The current model simplifies human intent, making it insufficient

to fully capture the complexity of human behavior. In reality, human decision-making process may be influenced by factors such as bounded rationality under time constraints, limited information, and risk awareness. As a result, actual human behavior may deviate from the robot's predicted optimal outcome [55], [56]. Thus, the incorporation of a more advanced human decision-making model may enhance the robot control strategy in more complex human–robot collaborative tasks.

VI. Conclusion

This article develops an empathetic intent inference framework through a noncooperative two-agent nonzero-sum game to address the double-blindness issue in HRC. A PNE control policy is proposed, enabling the robot to take proper actions to achieve effective collaboration. The results of two human subjects experiments empirically validate that the designed robot agent exhibits human-like behaviors and can be applied to complex collaborative tasks.

For future work, we plan to explore approaches to enabling the robot agent to incorporate the bounded rationality of the human in the decision-making process. An approach where the robot uses current prediction accuracy to adjust its confidence in future predictions holds promise [57]. Finally, we aim to apply the proposed method to real-world HRC tasks to further assess its practicality and effectiveness.

Acknowledgment

The authors would like to thank Dr. Yi Ren and Dr. Wenlong Zhang for providing the theoretical foundations of this work and they also would like to thank Jiayi Liu and Hainan Ding for their valuable help in building the experimental platform.

References

[1] Y. Sheng, H. Cheng, Y. Wang, H. Zhao, and H. Ding, "Teleoperated surgical robot with adaptive interactive control architecture for tissue identification," *Bioengineering*, vol. 10, no. 10, 2023, Art. no. 1157.

[2] R. Jahanmahin, S. Masoud, J. Rickli, and A. Djuric, "Human-robot interactions in manufacturing: A survey of human behavior modeling," *Robot. Comput.- Integr. Manuf.*, vol. 78, 2022, Art. no. 102404.

[3] Z. Zhou, C. Rother, and J. Chen, "Event-triggered model predictive control for autonomous vehicle path tracking: Validation using Carla simulator," *IEEE Trans. Intell. Veh.*, vol. 8, no. 6, pp. 3547–3555, Jun. 2023.

[4] Y. Li, K. P. Tee, S. S. Ge, and H. Li, "Building companionship through human-robot collaboration," in *Proc. Social Robot.: 5th Int. Conf.*, 2013, pp. 1–7.

[5] F. Semeraro, A. Griffiths, and A. Cangelosi, "Human–robot collaboration and machine learning: A systematic review of recent research," *Robot. Comput.- Integr. Manuf.*, vol. 79, 2023, Art. no. 102432.

[6] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," *Robot. Auton. Syst.*, vol. 42, no. 3/4, pp. 143–166, 2003.

[7] T. Fukuda, R. Michelini, V. Potkonjak, S. Tzafestas, K. Valavanis, and M. Vukobratovic, "How far away is" artificial man," *IEEE Robot. Autom. Mag.*, vol. 8, no. 1, pp. 66–73, Mar. 2001.

[8] M. J. Gielniak, C. K. Liu, and A. L. Thomaz, "Generating human-like motion for robots," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1275–1301, 2013.

[9] P. Hang, C. Lv, Y. Xing, C. Huang, and Z. Hu, "Human-like decision making for autonomous driving: A noncooperative game theoretic approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2076–2087, Apr. 2021.

[10] A. Takagi, G. Ganesh, T. Yoshioka, M. Kawato, and E. Burdet, "Physically interacting individuals estimate their partner's movement goal to enhance motor abilities," *Nature Hum. Behav.*, vol. 1, pp. 1–6, 2017.

[11] D. Aarno and D. Kragic, "Motion intention recognition in robot assisted applications," *Robot. Auton. Syst.*, vol. 56, no. 8, pp. 692–705, 2008.

[12] S. Nikolaidis, D. Hsu, and S. Srinivasa, "Human-robot mutual adaptation in collaborative tasks: Models and experiments," *Int. J. Robot. Res.*, vol. 36, no. 5–7, pp. 618–634, 2017.

[13] C. Liu, W. Zhang, and M. Tomizuka, "Who to blame? learning and control strategies with information asymmetry," in *Proc. Amer. Control Conf.*, IEEE, 2016, pp. 4859–4864.

[14] R. Tian, M. Tomizuka, A. D. Dragan, and A. Bajcsy, "Towards modeling and influencing the dynamics of human learning," in *Proc. ACM/IEEE Int. Conf. Hum.-Robot Interact.*, 2023, pp. 350–358.

[15] N. Rabinowitz, F. Perbet, F. Song, C. Zhang, S. A. Eslami, and M. Botvinick, "Machine theory of mind," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 4218–4227.

[16] C. Yu, B. Serhan, and A. Cangelosi, "Top-tom: Trust-aware robot policy with theory of mind," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2024, pp. 7888–7894.

[17] Y. Wang, Y. Ren, S. Elliott, and W. Zhang, "Enabling courteous vehicle interactions through game-based and dynamics-aware intent inference," *IEEE Trans. Intell. Veh.*, vol. 5, no. 2, pp. 217–228, Jun. 2020.

[18] A. J. Elliot and T. M. Thrash, "Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals," *J. Pers. social Psychol.*, vol. 82, no. 5, pp. 804–818, 2002.

[19] A. Kucukyilmaz, T. M. Sezgin, and C. Basdogan, "Conveying intentions through haptics in human-computer collaboration," in *Proc. IEEE World Haptics Conf.*, 2011, pp. 421–426.

[20] A. Gottardi, S. Tortora, E. Tosello, and E. Menegatti, "Shared control in robot teleoperation with improved potential fields," *IEEE Trans. Hum.-Mach. Syst.*, vol. 52, no. 3, pp. 410–422, Jun. 2022.

[21] A. Mörtl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The role of roles: Physical cooperation between humans and robots," *Int. J. Robot. Res.*, vol. 31, no. 13, pp. 1656–1674, 2012.

[22] Y. Li, K. P. Tee, W. L. Chan, R. Yan, Y. Chua, and D. K. Limbu, "Continuous role adaptation for human–robot shared control," *IEEE Trans. Robot.*, vol. 31, no. 3, pp. 672–681, Jun. 2015.

[23] Q. Wan, Y. Shi, X. Xiao, X. Li, and H. Mo, "Review of human–robot collaboration in robotic surgery," *Adv. Intell. Syst.*, vol. 7, 2024, Art. no. 2400319.

[24] Y. Ou and M. Tavakoli, "Sim-to-real surgical robot learning and autonomous planning for internal tissue points manipulation using reinforcement learning," *IEEE Robot. Autom. Lett.*, vol. 8, no. 5, pp. 2502–2509, May 2023.

[25] A. Buerkle, W. Eaton, N. Lohse, T. Bamber, and P. Ferreira, "EEG based arm movement intention recognition towards enhanced safety in symbiotic human-robot collaboration," *Robot. Comput.- Integr. Manuf.*, vol. 70, 2021, Art. no. 102137.

[26] T. Lenzi, S. M. M. De Rossi, N. Vitiello, and M. C. Carrozza, "Intention-based EMG control for powered exoskeletons," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 8, pp. 2180–2190, Aug. 2012.

[27] C. Cui, G.-B. Bian, Z.-G. Hou, J. Zhao, and H. Zhou, "A multimodal framework based on integration of cortical and muscular activities for decoding human intentions about lower limb motions," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 4, pp. 889–899, Aug. 2017.

[28] D. Zhang, L. Yao, K. Chen, S. Wang, X. Chang, and Y. Liu, "Making sense of spatio-temporal preserving representations for EEG-based human intention recognition," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3033–3044, Jul. 2020.

[29] S. Jain and B. Argall, "Probabilistic human intent recognition for shared autonomy in assistive robotics," *ACM Trans. Hum.-Robot Interact.*, vol. 9, no. 1, pp. 1–23, 2019.

[30] B. I. Ahmad, J. K. Murphy, P. M. Langdon, and S. J. Godsill, "Bayesian intent prediction in object tracking using bridging distributions," *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 215–227, Jan. 2018.

[31] X. Yu et al., "Bayesian estimation of human impedance and motion intention for human–robot collaboration," *IEEE Trans. Cybern.*, vol. 51, no. 4, pp. 1822–1834, Apr. 2021.

[32] A. Zyner, S. Worrall, and E. Nebot, "A recurrent neural network solution for predicting driver intention at unsignalized intersections," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1759–1764, Jul. 2018.

[33] W. Qi, S. E. Ovur, Z. Li, A. Marzullo, and R. Song, "Multi-sensor guided hand gesture recognition for a teleoperated robot using a recurrent neural network," *IEEE Robot. Autom. Lett.*, vol. 6, no. 3, pp. 6039–6045, Jul. 2021.

[34] Y. Li, K. P. Tee, R. Yan, W. L. Chan, and Y. Wu, "A framework of human–robot coordination based on game theory and policy iteration," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1408–1418, Dec. 2016.

[35] Y. Li, G. Carboni, F. Gonzalez, D. Campolo, and E. Burdet, "Differential game theory for versatile physical human–robot interaction," *Nature Mach. Intell.*, vol. 1, no. 1, pp. 36–43, 2019.

[36] L. Pezeshki, H. Sadeghian, M. Keshmiri, X. Chen, and S. Haddadin, "Cooperative assist-as-needed control for robotic rehabilitation: A two-player game approach," *IEEE Robot. Autom. Lett.*, vol. 8, no. 5, pp. 2852–2859, May 2023.

[37] L. Sun, W. Zhan, and M. Tomizuka, "Probabilistic prediction of interactive driving behavior via hierarchical inverse reinforcement learning," in *Proc. 21st Int. Conf. Intell. Transp. Syst.*, 2018, pp. 2111–2117.

[38] P. Chen, H. Zhao, X. Zhao, D. Ge, and H. Ding, "Dimensionality reduction for motion planning of dual-arm robots," in *Proc. IEEE Int. Conf. Mechatron. Autom.*, 2018, pp. 718–723.

[39] I. Raño and I. Iossifidis, "Modelling human arm motion through the attractor dynamics approach," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2013, pp. 2088–2093.

[40] C. Lauretti, F. Cordella, and L. Zollo, "A hybrid joint/cartesian DMP-based approach for obstacle avoidance of anthropomorphic assistive robots," *Int. J. Social Robot.*, vol. 11, no. 5, pp. 783–796, 2019.

[41] S. Kim, C. Kim, and J. H. Park, "Human-like arm motion generation for humanoid robots using motion capture database," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2006, pp. 3486–3491.

[42] C. Kim, S. Kim, S. Ra, and B. J. You, "Regenerating human-like arm motions of humanoid robots for a movable object," in *Proc. SICE Annu. Conf.*, 2007, pp. 1081–1086.

[43] M. Zare, P. M. Kebria, A. Khosravi, and S. Nahavandi, "A survey of imitation learning: Algorithms, recent developments, and challenges," *IEEE Trans. Cybern.*, vol. 54, no. 12, pp. 7173–7186, Dec. 2024.

[44] G. Xiang and J. Su, "Task-oriented deep reinforcement learning for robotic skill acquisition and control," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 1056–1069, Feb. 2021.

[45] M. Lindorfer, C. F. Mecklenbraeuker, and G. Ostermayer, "Modeling the imperfect driver: Incorporating human factors in a microscopic traffic model," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 9, pp. 2856–2870, Sep. 2018.

[46] A. Li, H. Jiang, Z. Li, J. Zhou, and X. Zhou, "Human-like trajectory planning on curved road: Learning from human drivers," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 8, pp. 3388–3397, Aug. 2020.

[47] L. Jiang and Y. Wang, "A personalized computational model for human-like automated decision-making," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 2, pp. 850–863, Apr. 2022.

[48] Y. Wang, P. Shintre, S. Amatya, and W. Zhang, "Bounded rational game-theoretical modeling of human joint actions with incomplete information," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2022, pp. 10720–10725.

[49] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Philadelphia, PA, USA: SIAM, 1998.

[50] Y. Jin, S. Yang, W. Lv, H. Yu, S. Zhu, and J. Huang, "Structure entropy minimization-based dynamic social interaction modeling for trajectory prediction," *Inf. Sci.*, vol. 614, pp. 170–184, 2022.

[51] F. Alambeigi, Z. Wang, Y.-h. Liu, R. H. Taylor, and M. Armand, "Toward semi-autonomous cryoablation of kidney tumors via model-independent deformable tissue manipulation technique," *Ann. Biomed. Eng.*, vol. 46, pp. 1650–1662, 2018.

[52] P. M. Scheikl et al., "Sim-to-real transfer for visual reinforcement learning of deformable object manipulation for robot-assisted surgery," *IEEE Robot. Autom. Lett.*, vol. 8, no. 2, pp. 560–567, Feb. 2023.

[53] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, Elsevier, 1988, pp. 139–183.

[54] Y. Wang, G. J. Lematta, C.-P. Hsiung, K. A. Rahm, E. K. Chiou, and W. Zhang, "Quantitative modeling and analysis of reliance in physical human–machine coordination," *J. Mech. Robot.*, vol. 11, no. 6, 2019, Art. no. 060901.

[55] S. Schach, S. Gottwald, and D. A. Braun, "Quantifying motor task performance by bounded rational decision theory," *Front. Neurosci.*, vol. 12, pp. 932–951, 2018.

[56] M. Kwon, E. Biyik, A. Talati, K. Bhasin, D. P. Losey, and D. Sadigh, "When humans aren't optimal: Robots that collaborate with risk-aware humans," in *Proc. ACM/IEEE Int. Conf. Hum.-Robot Interact.*, 2020, pp. 43–52.

[57] D. Fridovich-Keil et al., "Confidence-aware motion prediction for real-time collision avoidance[1]," *Int. J. Robot. Res.*, vol. 39, no. 2/3, pp. 250–265, 2020.

Yubo Sheng (Graduate Student Member, IEEE) received the B.Eng. degree in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2017 and the M.Sci. degree in mechanical engineering from the HuaZhong University of Science Technology, Wuhan, China, in 2022. He is currently working toward the Ph.D. degree in mechanical engineering at Huazhong University of Science and Technology.

He conducted research at the State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology. His research includes human–robot interaction and telerobotic systems.

Yiwei Wang (Member, IEEE) received the B.Eng. degree in mechanical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2013 and the M.Sci. and Ph.D. degrees in mechanical engineering from Arizona State University, Tempe, AZ, USA, in 2014 and 2021, respectively.

He was a Postdoctoral Researcher with State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology. He is currently an Assistant Researcher with Institute of Medical Equipment Science and Engineering, Huazhong University of Science and Technology. His research interests include physical human–robot interaction, human–robot collaboration, and surgical robots.

Haoyuan Cheng received the B.Eng. degree in mechanical engineering from the Wuhan University of Technology, Wuhan, China, in 2021. He is currently working toward the M.Sci. degree in mechanical engineering with the Huazhong University of Science and Technology, Wuhan, in 2024.

He conducted research at the State Key Laboratory of Intelligent Manufacturing Equipment and Technology, Huazhong University of Science and Technology. His research focuses on minimally invasive surgical robots and teleoperation.

Huan Zhao (Member, IEEE) received the B.E. degree in mechanical engineering from the School of Mechanical Science and Engineering, Jilin University, Changchun, China, in 2006 and the Ph.D. degree in mechanical engineering from the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2013.

From 2013 to 2015, he was a Postdoctoral Researcher with the Huazhong University of Science and Technology, Wuhan, China. Since 2015, he has been with the Huazhong University of Science and Technology, where he is currently a Professor. His research interests include robotic machining and assembly.

Han Ding (Senior Member, IEEE) received the Ph.D. degree from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1989.

Supported by the Alexander von Humboldt Foundation, he was with the University of Stuttgart, Stuttgart, Germany, from 1993 to 1994. Since 1997, he has been a Professor with HUST, where he is currently the Dean of the Future Institute of Technology. He was a Cheung Kong Chair Professor with Shanghai Jiao Tong University, Shanghai, China, from 2001 to 2006. He was a member of the Chinese Academy of Sciences in 2013. His research interests include robotics, control engineering, and digital manufacturing.

Dr. Ding is currently an Editor of IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING and a Senior Editor of IEEE ROBOTICS AND AUTOMATION LETTERS.