Combining Learning from Demonstration with Learning by Exploration to Facilitate Contact-Rich Tasks

Yunlei Shi[1,2], Zhaopeng Chen[2,1], Yansong Wu[3], Dimitri Henkel[2], Sebastian Riedel[2], Hongxu Liu[3], Qian Feng[3,2], Jianwei Zhang[1]

***Abstract*— Collaborative robots are expected to be able to work alongside humans and in some cases directly replace existing human workers, thus effectively responding to rapid assembly line changes. Current methods for programming contact-rich tasks, especially in heavily constrained space, tend to be fairly inefficient. Therefore, faster and more intuitive approaches to robot teaching are urgently required. This work focuses on combining visual servoing based learning from demonstration (LfD) and force-based learning by exploration (LbE), to enable fast and intuitive programming of contact-rich tasks with minimal user effort required. Two learning approaches were developed and integrated into a framework, and one relying on human to robot motion mapping (the visual servoing approach) and one on force-based reinforcement learning. The developed framework implements the non-contact demonstration teaching method based on visual servoing approach and optimizes the demonstrated robot target positions according to the detected contact state. The framework has been compared with two most commonly used baseline techniques, pendant-based teaching and hand-guiding teaching. The efficiency and reliability of the framework have been validated through comparison experiments involving the teaching and execution of contact-rich tasks. The framework proposed in this paper has performed the best in terms of teaching time, execution success rate, risk of damage, and ease of use.**

I. INTRODUCTION

The ability to rapidly setup and reprogram newly-introduced products is an increasingly essential requirement for adaptive robotic assembly systems in factories [1], [2]. Position-controlled robots can handle known objects within well-structured assembly lines with high efficiency and achieve highly accurate position control. However, they require significant time to set up and tedious reprogramming to fulfill new tasks, and cannot adapt to any unexpected variations in assembly processes [3].

Collaborative robots hold the promise of closing the onerous reprogramming and unexpected variations gap by combining the capabilities of position-controlled robots with dexterity and flexibility. For example, the hand-guiding method enables unskilled users to interact with collaborative robots and program quickly [4]. However, when the assembly line needs to be reconfigured, it still takes a long time to remove and reinstall the robot arms and various attachments.

*This research has received funding from the German Research Foundation (DFG) and the National Science Foundation of China (NSFC) in project Crossmodal Learning, DFG TRR-169/NSFC 61621136008, partially supported by European projects H2020 STEP2DYNA (691154) and ULTRACEPT (778602).

[1]TAMS (Technical Aspects of Multimodal Systems), Department of Informatics, Universität Hamburg, [2]Agile Robots AG, [3]Technische Universität München.

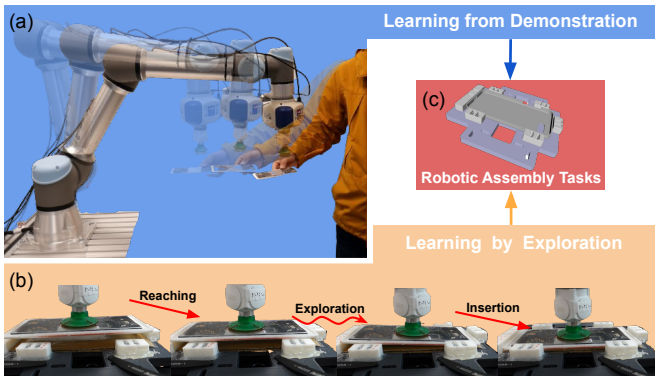

Fig. 1. The robot arm and suction gripper performing a contact-rich tending task. (a) The gross motion is learned from the human's demonstration. (b) The fine motion is learned from exploration. (c) An example of contact-rich tending task

Mobile manipulators (where robotic arms are mounted on mobile bases) were introduced to expand the productivity and adaptive capacity of manufacturing automation especially for the setting up phase when production lines must be reconfigured [5], [6]. Since mobile manipulators can only be placed beside the production lines instead of installing on the production line like collaborative robots, which occupying more space previously provided for human workers. However, programming a robot in a constrained space is very difficult [7]. Overall, ease-of-programming has been identified as an open challenge in robot assembly [2], [8].

Additionally, collaborative robots equipped with force control functions can perform certain hybrid position/force operations for contact-rich tasks [9], [10], [11], [12], [13], while their effectiveness and variation adaptive capacity in assembly scenarios are still not satisfactory [14], [15].

Herein, we propose an intuitive programming method to decrease the setting up time of a mobile manipulator, and introduce a reinforcement learning (RL) algorithm to overcome unexpected variations in assembly tasks as shown in Figure 1. Our primary contributions are:

- **C1:** A approach that learns the trajectory from demonstrations based on visual servoing for fast, easy, accurate setting up in heavily constrained space scenarios.
- **C2:** A region-limited residual reinforcement learning (RRRL) policy based on force-torque information was trained to overcome the pose uncertainty.
- **C3: C1** and **C2** were combined into a robot tending skill and was compared with two most commonly used baseline techniques using a UR5e robot.

The paper is organized as follows: in Section II, we provide an overview of the development of industrial mobile manipulators and programming methods, as well as their pros and cons. Moreover, the force controller is briefly introduced. In Section III, we outline some of the key problems and corresponding ideas. Details on our proposed method are provided in section IV. A quantitative experiment of our methods is presented in Section V. Finally, we present and discuss the experimental results in Section VI, and suggest some possible areas of future research in Section VII.

II. RELATED WORK AND BACKGROUND

A. Industrial Mobile Manipulator

The mobile manipulator system (also known as mobile manipulator hybrid robots[1] or collaborative cells[2]) was designed to reduce the setup time [16]; all required devices of the tasks are integrated into the mobile platform and it therefore a "plug and play" solution. Various mobile manipulators have been developed, such as Rob@Work [17], Little Helper [18], KMR iiwa [19], KUKA flexFELLOW[3] and MiR UR[4], which are the most widely-deployed industrial mobile manipulators. However, the traditional magnetic tracking approach utilized by commercial mobile manipulators (i.e., automatic guided vehicle (AGVs)) can be accompanied by the position uncertainty problem, which is incompatible with the requirement of many industrial applications, especially when the mobile manipulators must be moved between picking and placing locations.

B. LfD Under Visual Guidance

LfD has been recently suggested as an effective means of speeding up the programming of learning processes, spanning low-level control to high-level assembly planning [20]. Guiding robots by means of visual feedback [21], [22], [23] during assembly tasks is an effective way to overcome the position uncertainties introduced by mobile manipulators. For robotic assembly tasks, executing high precision measurements is important. However, visual errors could be introduced by lens and the imaging sensors, as well as the calibration of intrinsic and extrinsic parameters [24]. Some researchers hold the view that humans should focus more on the execution tasks than the vision sensors. In accordance with this notion, other approaches have been considered, such as intelligent assembly algorithms, in an effort to lower the necessity of vision sensors for given tasks.

C. LbE for Contact-rich Tasks

LbE has also been suggested as an effective means of reducing programming time, and recent researches have introduced artificial intelligence methods into robotics [1], [3], [25], [26]. RL offers a set of tools for designing sophisticated robotic behaviors that are difficult to engineer.

[1]http://std.samr.gov.cn/
[2]https://blog.robotiq.com/
[3]www.kuka.com
[4]https://mobile-industrial-robots.com/
[5]www.kuka.com/en-de/products/robot-systems/industrial-robots/lbr-iiwa

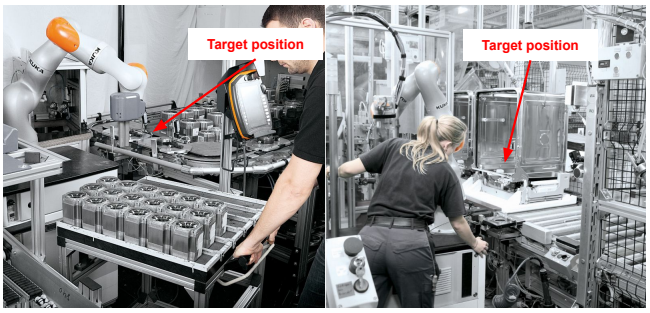

Fig. 2. Examples of heavily constrained space in real factories [5].

RL and its derivative methods have previously been applied with great success to solve a variety of robotic manipulation problems [13], [27], [10], [14], [28]. Exploration entails interaction between the robot and the operational environment; therefore, a force or impedance controller is required.

D. Operational Space Force Control

Interaction control attracted widespread scholarly interest following the pioneering work on impedance/admittance control and passivity by Hogan [29] and Colgate [30]. However, the first mentions of admittance control concepts date back to Whitney [31], where they were used to address hard contact issue for industrial manipulation functions and indirect force control purposes. Overall, operational space force control has been proven the effectiveness of RL-based assembly tasks [10], [14].

III. PROBLEM STATEMENT AND METHOD OVERVIEW

A. Programming with Mobile Manipulator

1) Problem details: Mobile manipulators can significantly reduce installation times [32]. Programming based on demonstration approaches has been proposed to address variations in geometry and configurations for assembly, placement, handling, and picking manipulations [32], which could reduce both programming time and user training requirements [3]. The use of mobile manipulators introduces a **positioning error at the $\pm$5 mm level** [33], and errors as small as $\pm$1 mm can lead to **huge contact forces and consistent failures** among typical assembly tasks [15]. Whereas in this study we choose to address the more typical cases of the mobile bases of mobile manipulators being repositioned in accordance with task requirements.

Teach pendants are still in use for precision positioning (position and orientation of the end effector (EE) for many tasks [4]. However, these devices not only limit the intuitiveness of the teaching process but are also time-consuming to use. Hand-guiding is a typical physical contact kinesthetic teaching solution, which embodies programming through demonstration concepts, enabling users to program robots quickly and intuitively. However, it has drawbacks in terms of accuracy, large locational separations, and operations involving dangerous objects [3]. More seriously, neither the hand-guiding nor the teach pendant programming method can compensate for the major positioning error that accompanies

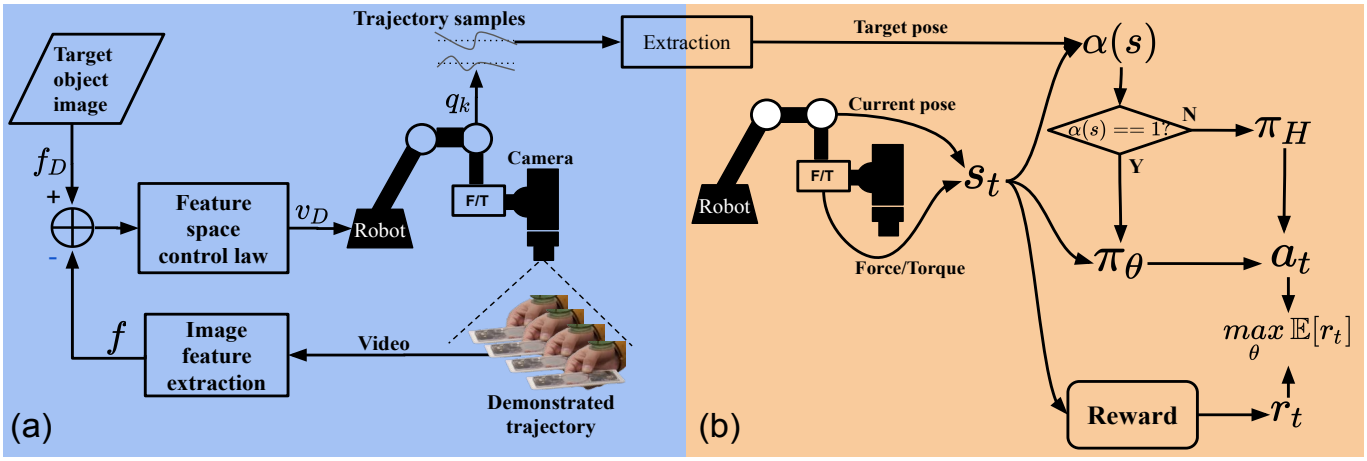

Fig. 3. Combination of LfD policy (a) and RRRL policy (b)

mobile units [33], and can result in the generation of huge contact force that can damage objects [4].

In addition, mobile manipulator's base takes up significant space within work cells, with Figure 2 displaying some real-world factory examples of this. Sometimes, a user must teach the robot with a highly **awkward body posture** due to anthropometric limitations [7]; moreover, the **resistance of robots in drag mode** makes the user feel difficult to move delicately [7], these two issues make the hand-guiding for accurate assembly rather difficult and low quality (e.g., excessive contact force and low accuracy). Robots equipped with joint torque sensors partially solved the second issue of resistance, while the UR5e used in our experiment is widely used in the industry, and many users would still be affected.

2) Method overview: We proposed a method that is simple and fast to implement and reduce the physical contact force to solve the aforementioned problems in Section III-A. We introduce visual guiding into the teaching phase as shown in Figure 3(a), our visual guiding teaching is performed with 2 steps:

- **Grasping pose definition:** First, the user guides the robot to define a grasping pose under the target object frame. Then the robot moves up and uses the eye-in-hand camera to take a photo of the object as a reference.
- **Trajectory generation:** Second, the robot follows the moving object (the object could be moved by user) to the new target pose using a vision-based control algorithm (e.g., visual servoing [34]), and records whole moving trajectory.

Our method performs better in the real industrial scenario than the report by Wei et al [35] because they used a global camera that could be occluded easily while we use an eye-in-hand camera and kept the robot following the object to achieve a proper trajectory. The trajectory error has little effect on the final assembly because uncertainties can be generally ignored in gross motion planning and a fine motion planner will solve the uncertainties during the assembly process [36].

B. Assembly Under Position Uncertainty

1) Problem details: The visual sensor can be used for target recognition, pose estimation, measuring, and positioning using the traditional methods [24]. While for visual sensors, not only the lens and imaging sensors, the calibration of intrinsic/extrinsic parameters will exert a strong influence on the precision of the visual guidance, but also the reflection, shadow, and occlusion may fail the extraction of the object edge or features owing to the change of light and the texture of objects [24].

2) Method overview: Residual reinforcement learning (Residual RL) [15], [37] is a novel method for taking advantage of the conventional controllers'efficiency and the RL's flexibility. Residual RL attempts to inject prior information into an RL algorithm to accelerate the training process instead of randomly exploring from scratch. For example, the estimated position can be set as prior information and even it could have errors. Guided Uncertainty Aware Policy Optimization (GUAPO) [28] showed better performance than residual RL, SAC, and other pure model-based methods. However, the force and torque information was not carried in the GUAPO policy. The force and torque information could provide observations of current contact conditions between the object and environment for an accurate localization [21] and ensure the safety of manipulation [14], however,pure force-based learning policy may lead to substantial deviation from the goal and reduce the learning efficiency. Thus, we combine the "region limitation" idea from GUAPO and residual RL policy to develop a force-based approach called "**R**egion-limited **R**esidual **RL**" (**RRRL**) as shown in Figure 3(b). In the RRRL policy, the rough target pose is supplied by the teaching phase as the residual part, a function $\alpha(s) = 1[s \in \mathbb{S}_u]$ is introduced to switch between the fixed policy $\pi_H(s)$ and the parametric policy $\pi_\theta(s)$ [28]:

$$\pi(a|s) = (1 - \alpha(s)) \cdot \pi_H(a|s) + \alpha(s) \cdot \pi_\theta(a|s). \quad (1)$$

$\mathbb{S}_u$ is the region that containing the goal position with uncertainty. Since force control is more safe and reliable in the fine motion/manipulation phase than position control and impedance control in assembly tasks [10], the RRRL policy

$\pi_\theta(s)$ takes the operational force controller as the desired force/torque in operational space; our goal is to optimize them through the RRRL policy. The fixed policy u_H is used to move the object back to the initial target pose when the function $\alpha(s) = 0$.

IV. POLICY DESIGN

The processes in LfD could be divided into three steps: observe an operation, represent the operation and reproduce the operation [3]. We perform our whole policy similar to the aforementioned processes, and add the RRRL policy at the end of the operations. Our method consists of two learning policies:

- **Lfd policy**: the robot learns gross motion via human's demonstration, whereby a human will demonstrate the object image and grasping position to the robot.
- **RRRL Policy**: the robot learns fine motion based on the RRRL policy, as described in Algorithm 1. The RRRL policy can be trained in advance in order to save the setup time.

First of all, we define the terminology and notation required in order to represent the coordinate transformations. We represent the task space of the robot as $\mathcal{T}$, which constitutes the set of positions and orientations that the robotic EE (i.e., suction gripper) can attain. $\mathcal{T}$ is a smooth m-manifold in which m = 6 and $\mathcal{T} = SE^3 = \mathcal{R}^3 \times SO^3$. The superscripts/subscripts of the coordinate frames that will be used are listed in Table I.

TABLE I
COORDINATE FRAMES

e	The coordinate frame attached to the robot EE
c	The coordinate frame of the camera
b	The base coordinate frame of the robot
o	The coordinate frame attached to the target object

A. LfD Policy

In our policy, we equip an eye-in-hand camera in order to avoid the occlusion caused by the robot's links and other industrial devices [3], [38] during the demonstration. The relative homogeneous transformation ${}^e\boldsymbol{x}_c$ and intrinsic camera parameters were determined by means of the hand/eye calibration method [39].

During the demonstration phase, the teaching was separated into the three following steps:
1) Robot EE was moved to the **desired grasping pose (DGP)**, which was then recorded as ${}^b\boldsymbol{x}_{DGP}$.
2) Robot EE was moved to the **desired visual servoing pose (DVSP)**, which was recorded as ${}^b\boldsymbol{x}_{DVSP}$, and here the object needed to be kept in view of the camera. The **first reference photo (RF1)** was taken, together we calculated a fixed **relative pose (RP)**:

$$ {}^c\boldsymbol{x}_o = {}^{DVSP}\boldsymbol{x}_{DGP} = ({}^b\boldsymbol{x}_{DVSP})^{-1}({}^b\boldsymbol{x}_{DGP}), \quad (2) $$

${}^c\boldsymbol{x}_o$ is the coordinate transformation of object frame o respect to camera frame c.

Algorithm 1 RRRL

Require: Model based policy π_H, learning frequency C_1, target action-value update frequency C_2.

```
Initialize replay memory H to capacity N
Initialize action-value function Q with random weights θ
Initialize target action-value function Q_target with weights θ⁻ = θ
for episode = 1 to M do
    Sample state s_0
    while NOT EpisodeEnd do
        Calculate α(s) with Equation (8)
        Choose action a_H from π_H(s_t)
        With probability ϵ choose a random action a_RL
        Otherwise select a_RL ∼ π_θ(s_t)
        Obtain action a_t = (1 − α) ∗ a_H + α ∗ a_RL
        Execute a_t, observe reward r_t and state s_{t+1}
        Store transition (s_t, a_t, r_t, s_{t+1}) in H with priority p_t = max_{i<t} p_i
        for j = 1 to C_1 do
            Sample minibatch of transitions with priority from H
            Update transition priority
            Update θ with the method proposed in [40]
        end for
        Every C_2 steps reset Q_target = Q
    end while
end for
```

3) The visual servoing strategy was activated, and the following system constraints [41] were applied to the robot during the teaching process [42]:

$$ \mathbf{q} \in \mathbb{Q}_c, \quad (3) $$

$$ \mathbf{q} \in [\mathbf{q}^{min}, \mathbf{q}^{max}], \quad \mathbf{q}^{min}, \mathbf{q}^{max} \in \mathbb{R}^N, \quad (4) $$

$$ \dot{\mathbf{q}} \in [\dot{\mathbf{q}}^{min}, \dot{\mathbf{q}}^{max}], \quad \dot{\mathbf{q}}^{min}, \dot{\mathbf{q}}^{max} \in \mathbb{R}^N, \quad (5) $$

$$ \mathbf{q}_{k+1} = \mathbf{q}_k + \delta\dot{\mathbf{q}}_k, \quad (6) $$

$\mathbb{Q}_c$ is the set of configurations that do not cause any part of the arm to collide with obstacles that are difficult to model. Equation (4) and Equation (5) describe the robots joint position and velocity constraints, respectively. The object was then moved from the **DVSP** to the **desired final pose (DFP)** by the user, and the robot EE will follow the trajectory $\mathbf{q}_k$ from the **DVSP** to the **DFP** under the constraints, the trajectory could be recorded. At **DFP**, the camera takes the **second reference photo (RF2)** automatically, and the **DFP** could be calculate easily at the end of the trajectory by Equation (7):

$$ DFP = ({}^b\boldsymbol{x}_e)({}^e\boldsymbol{x}_c)({}^c\boldsymbol{x}_o) \quad (7) $$

In this paper, the image-based visual servoing (IBVS) method was introduced based on the specified observed feature positions as shown in Figure 3(a). With the human expert (user) in the teaching loop, trajectory-based representation

output by LfD could avoid the collision even if the robot operates in a heavily constrained operation space. Besides, our non-contact guiding method could avoid the resistance force of the robot in drag mode which makes the user perform the teaching more easily and more accurately.

B. Region-limited Reinforcement Learning Policy (RRRL)

In this section, we explain the RRRL algorithm details as shown in Figure 3(b) and Algorithm 1 for learning the assembly tasks.

1) Limited Region and Policy Structure: Most assembly tasks can be characterized as minimizing the distance between objects and their goal positions [14]. The limited region is used to constrain the area of exploration. Many methods can describe the uncertainty region by reference to a nonparametric distribution [28] or parametric equation. In this paper, a Euclidean distance was used to indicate the switch signal between the fixed policy $\pi_H(s)$ and parametric policy $\pi_\theta(s)$ in the hybrid RRRL policy:

$$\alpha(s) = \begin{cases} 1, & \text{if } \|CP - DFP\|_2 < D \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

CP and DFP are the current pose and desired final pose, respectively. D is an engineering hyperparameter that is determined on the basis of experience, and we suggest that D should be at least twice as large as the final positioning error introduced by the fixed policy $\pi_H(s)$. When $\alpha(s) = 1$, the **force-based** learning policy $\pi_\theta(s)$ will be called upon, otherwise the **position-based** fixed policy $\pi_H(s)$ will bring the object back to its initial uncertain target pose. The same switch algorithm is used during the training and execution phase. A double DQN with proportional prioritization [40] was selected as the learning policy $\pi_\theta(s)$ in this study.

2) Action Design: The actions in the assembly task could be either position command [21], [28] or force/torque command [10], [14]. As we sought to reduce the contact force between the object and environment to ensure safety, the force/torque command action in the operational space (i.e. under frame x_e) was chosen [14].

Carefully taking advantage of the natural constraints in the design of learning policy is essential for the assembly tasks considered herein. It is obvious that the task simplifies, if the motion is constrained in "wrong" directions. We utilized comparative experiments to investigate how different force-based actions utilize natural constraints. We set the same initial positional error of $\delta P \in [2, 4]$ mm in a random direction, and then performed a random strategy to select the actions. We tested each action 200 times with a maximum of 20 steps and a maximum force amplitude of 10N, the success rates of which were as follows:

	Force control actions	Success rate
1	Operational space controller [14]	32%
2	Fz with the force of another 1 dimension [10]	60%
3	Fz with the forces of other 2 dimensions	69%

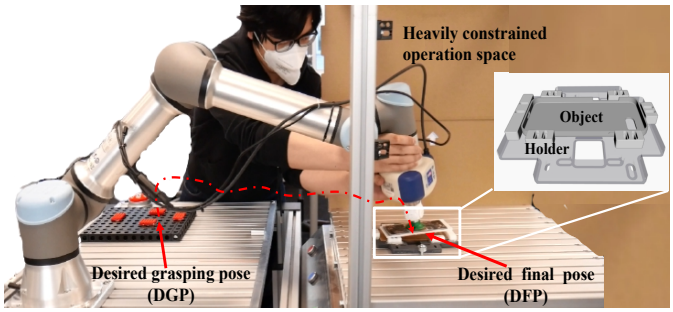

Fig. 4. Hand-guiding teaching in a heavily constrained operation space. Both the user and robot arms are close to their operating boundaries (near singular configurations).

The third discrete action as Equation (9) performs the best in the comparison experiment:

$$\begin{aligned} 1 &: [+{}^eF_x, +{}^eF_y, +{}^eF_z, 0, 0, 0] \\ 2 &: [+{}^eF_x, -{}^eF_y, +{}^eF_z, 0, 0, 0] \\ 3 &: [-{}^eF_x, +{}^eF_y, +{}^eF_z, 0, 0, 0] \\ 4 &: [-{}^eF_x, -{}^eF_y, +{}^eF_z, 0, 0, 0] \end{aligned} \tag{9}$$

Here, we set the orientation space as the position mode in order to take advantage of the flexibility of the suction cup; then, all of the torque commands in the operation space were set to 0 Nm and all the force amplitudes are set to 10 N, which is half of [10].

3) State Design: Forces and torques featured the most direct information that characterizes the contact states during operation, and so the 6-dimensional force-torque vector $s = [F_x, F_y, F_z, M_x, M_y, M_z]$ under the robot EE's frame x_e were sent to the RRRL networks as the input state. The state sampling frequency was 10 Hz, which could guarantee observation of the contact states.

4) Reward Design: We employed the precise target position of the hole as the reference for the reward during the learning phase. Unlike the execution phase, the precise desired position was easy to obtain, as the robot had high positional repeatability.

$$r = \begin{cases} 1 - k_{steps}/k_{max}, & \text{success} \\ -\|CP - DFP\|_2, & \text{otherwise.} \end{cases} \tag{10}$$

5) Parameters Setup: In order to gain more contact experience, a positional error $\delta P \in [2, 4]$ mm was added in a random direction during the training phase. The transitions (s_t, a_t, r_t, s_{t+1}) sampled from the environment was stored in a replay buffer [40]. The size of the experience's replay memory P_{replay} was 20, 000, the maximum number of training episodes M was set to 200, and the maximum number of steps k_{max} for the search phase was 50. The batch size P_{batch} was set as 64 in order to select random experiences from P_{replay}, and the discount factor λ was 0.5.

V. EXPERIMENTS

We evaluated our methods in **teaching** and **execution** phases. In the experiments in both phases, we aimed to answer the following questions: (**Q1**) Can our method maintain the fast and easy programming ability even in constrained

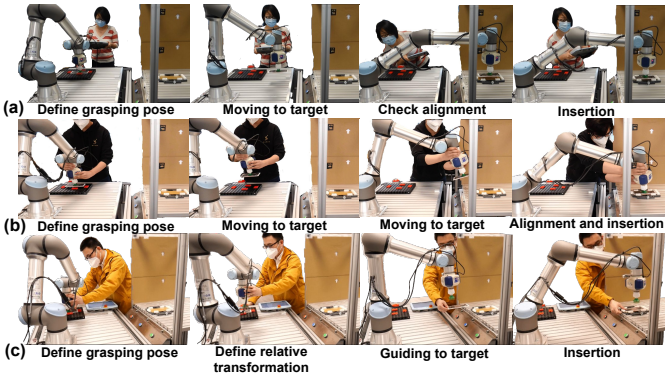

Fig. 5. (a) Teach-pendant teaching: it is difficult to align the object with the target holder by eye. (b) hand-guiding teaching:it is difficult to move the robot EE with singular configurations, also not easy to align. (c) our LfD method teaching: contact-free guiding that is not physically demanding, and is easy to align without robot resistance.

operational space? (**Q2**) Can our method retain the execution success rate against positional uncertainty? (**Q3**) Can our method reduce the risk of damage during operation of the object?

A UR5e robot Universal Robots [5] was used to implement our novel approach in a tending task. The UR5e features a 6-axis and 5 kg payload, a working radius of 850 mm, and is equipped with a 6 DOF force/torque sensor on the EE. A Schmalz CobotPump ECBPi suction cup was installed beyond the force/torque sensor in order to ensure the detection of the contact force with the environment. An Intel RealSense Depth Camera D435i also attached at the EE in order to conduct the visual servoing process. Our policy was run on CPU of a Dell Precision 5510 laptop, and sent the updated position to the UR5e controller; we used ur-rtde [6] as the Python interface for controlling and receiving data from the UR robot. A 6 DoF ATI Axia80 force sensor [7] was mounted under the holder in order to measure the operating force, and a low-pass filter with 9.37Hz cutoff frequency was used to mitigate the force noise.

A. Baseline Techniques

The commonly available methods used for a collaborative robot were selected as baselines [7]. As such, we compared our proposed method to the following baselines:

Baseline 1: Teach-pendant + spiral searching. The UR5e teaching pendant with a UR PolyScope GUI was held by users in one hand, who pressed the on-screen buttons to map the rate control of the EEs translation and rotation in the task space $\mathcal{T}$ with the other. At the DFP, a spiral search function similar to that outlined in [43] was added.

Baseline 2: Hand-guiding + spiral searching. The UR5e "Freedrive" mode was utilized, and users physically grabbed and exerted force to move the robot arm using one or two hands. At the DFP, a spiral search function was also added.

[5]https://www.universal-robots.com/products/ur5-robot/
[6]https://pypi.org/project/ur-rtde/
[7]https://www.ati-ia.com/Products/ft/sensors.aspx

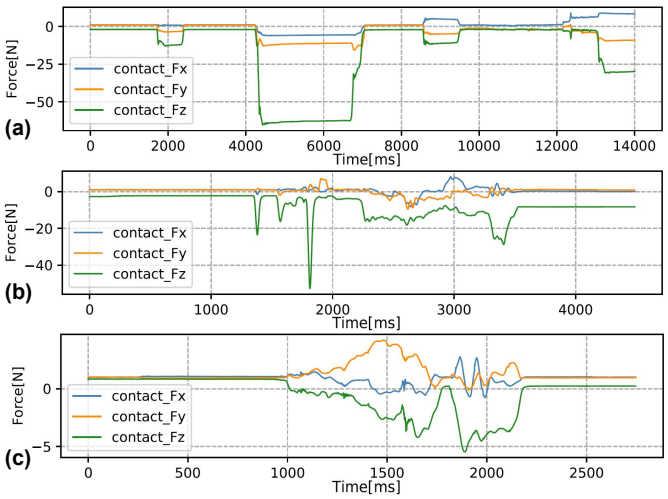

Fig. 6. Contact force during teaching. (a) This curve shows the contact characteristics of the Teach-pendant teaching method: the user uses an "observe-move" strategy and adjusts only after finding unsuccessful insertion, so the contact force will be maintained during observation. (b) The hand-guiding method: the object frequently collides with the holder when it is not inserted (impact force in the figure), and there is a continuous contact force after inserting the object. (c) Our method produces a small contact force (less than 5N) in the teaching phase.

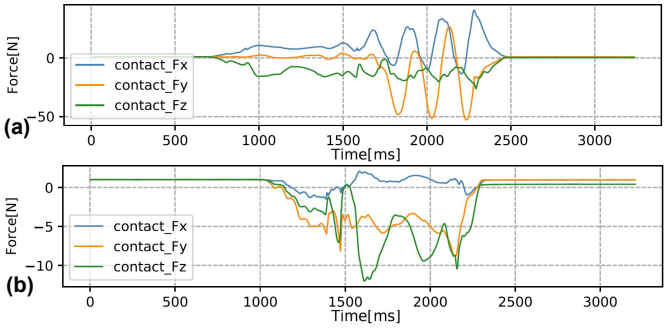

Fig. 7. Contact force during execution. (a) Spiral exploration method: the first half of the curve (700ms-1500ms) is not yet constrained by the holder, the contact force is small and stable, after 1600ms the object is inserted into the holder, immediately generating a larger contact force, due to the UR5e force sensor accuracy issue, the spiral movement generates a larger contact force than the stopping threshold. (b) RRRL method: the maximum contact force is around 10N because of the limitation of force action amplitude.

B. Task Setup

A tight clearance machine tending task (similar to the assembly task) was used to evaluate our method and the two baselines as shown in Figure 4 . One holder was installed in an opaque box in order to simulate a situation in which the field of view in the production line is obscured and an object must be inserted into the holder. The user was given a brief tutorial and allowed to practice until s/he felt ready. A group of four able-bodied volunteers (1 female, 3 male, age: 24 to 35) participated.

C. Measurement and Evaluation

- **Teaching time:** The task's completion time was measured as the time it took the user to move the robot from the DGP to the DFP.
- **Execution success rate:** After the teaching phase was complete, the trajectory of the demonstration was exe-

TABLE II

EVALUATION IN THE TEACHING PHASE

Teaching phase	Time cost	Maximum contact force
Teach-pendant	60–120 s	15–50 N
Hand-guiding	**15–42 s**	30–60 N
Our method	**23–30 s**	**3–10 N**

TABLE III

EVALUATION IN THE EXECUTION PHASE

Execution phase	**Success rate**		**Maximum contact force**
	Perfect	**Uncertainty**	
Only teach-pendant	55/100	17/100	15 N
Only hand-guiding	33/100	5/100	15 N
Teach-pendant + spiral searching	69/100	47/100	35 N
Hand-guiding + spiral searching	51/100	33/100	35 N
Our method	**95/100**	**91/100**	**15 N**

cuted and the insertion success rate was tested. Contrary to the **"perfect"** group, an error of $\delta P \in [2, 4]$ mm in a random direction was added on DFP to simulate the pose uncertainties in the **"uncertainty"** group.

- **Risk of damage:** The risk of damage caused by the operating force has been ignored in previous studies [13], [27], [14], [28], [43]. In this study, the maximum absolute contact force during contact operation was employed to evaluate the risk of damage.

VI. RESULTS AND DISCUSSIONS

In total, 12 group robot teaching results and 1000 group robot execution results were recorded.

The teaching phase test scenarios can be seen in Figure 5, and the results are displayed in as Table II. Our method features a similar time cost to the hand-guiding baseline, but it is more generalized. With the hand-guiding method, male volunteers always required less teaching time than female ones due to the physically demanding. All volunteers noted that the robot "required too much force to move especially near the boundaries". Our method does not require physical contact with the robot, and therefore it is not physically demanding. All volunteers required significantly longer teaching times when the teach-pendant was used. A lot of time was spent aligning the object with the target, and the tricky sight angle made this even more difficult. In contrast, our method does not require a human to align the object to the target, but enables the robot arm to actively track the object in order to attain the target position, and therefore the setup is fast. Hence, the answer to **Q1** is yes. Furthermore, our method produces minimal contact force on the environment as shown in Figure 6(c), as our schematic approach is identical to the human tending one. The other two methods result in much less transparency during interactions with the environment due to the resistance of the robot arm itself or the inability to interact with the environment in terms of force.

The execution phase results can be seen in Table III; the success rate of our method far exceeded that of the other reference baselines. We found that **the elasticity of the suction cup** had a major influence on the accuracy of the demonstrated target position. The contact force at the end of the demonstration could induce **deformation of the suction cup** and so affect the actual target position. With our method, the robot EE did not contact the target environment at the end of the teaching phase, and so there was no contact force and therefore no deformation of the suction cup. Finally, high target position accuracy was achieved. In addition, the RRRL policy could determine force actions based on the contact state of the object and holder, and so greatly improving the insertion success rate. Our method could guarantee small contact forces due to the amplitude limitation of the force actions as shown in fig. 7(b). Therefore, the answers to **Q2** and **Q3** are also yes.

VII. CONCLUSIONS

Collaborative robots are rapidly making advances into new work environments, but there is a control and teaching bottleneck in the programming of contact-rich tasks.

In this study, we combined the visual servoing based LfD and force-based LbE to facilitate the rapid and intuitive execution of assembly tasks requiring minimal user expertise, involvement, and physical exertion. The efficiency of the methods was validated through a series of experiments entailing the execution of a tending task using a robot arm and suction cup system.

In an experiment that compared our method against two commonly used baselines, namely teach-pendant and hand-guiding teaching, in a challenging setting designed to simulate a heavily constrained operation space, which is very common in actual factories, we found that our method received the best feedback in terms of both subjective and objective evaluations.

We noted the negative effect of the elasticity of the suction cups on the accuracy of the position demonstrations, and in future hope to investigate how to take advantage of this elasticity with a learning approach. We also plan to analyze more contact-rich tending tasks and more types of grippers in order to refine our method and improve its generalizability.

ACKNOWLEDGEMENT

We would like to thank Chengjie Yuan, Chunyang Chang, Jingjing Sun, and Alexander Daimer for their contributions as volunteers and for their valuable feedback during the discussion of the experimental protocol. We also thank mechanical engineer Ningxin Lu for building the hardware mockup system.

REFERENCES

[1] A. Kramberger, B. Nemec, M. Gams, A. Ude *et al.*, "Learning of assembly constraints by demonstration and active exploration," *Industrial Robot: An International Journal*, 2016.

[2] C. Sloth, A. Kramberger, and I. Iturrate, "Towards easy setup of robotic assembly tasks," *Advanced Robotics*, vol. 34, no. 7-8, pp. 499–513, 2020.

[3] Z. Zhu and H. Hu, "Robot learning from demonstration in robotic assembly: A survey," *Robotics*, vol. 7, no. 2, p. 17, 2018.

[4] M. Safeea, R. Bearee, and P. Neto, "End-effector precise hand-guiding for collaborative robots," in *Iberian Robotics conference*. Springer, 2017, pp. 595–605.

[5] J. Marvel and R. Bostelman, "Towards mobile manipulator safety standards," in *2013 IEEE International Symposium on Robotic and Sensors Environments (ROSE)*. IEEE, 2013, pp. 31–36.

[6] M. Wojtynek, J. J. Steil, and S. Wrede, "Plug, plan and produce as enabler for easy workcell setup and collaborative robot programming in smart factories," *KI-Künstliche Intelligenz*, vol. 33, no. 2, pp. 151–161, 2019.

[7] J. H. Lee, Y. Kim, S.-G. An, and S.-H. Bae, "Robot telekinesis: application of a unimanual and bimanual object manipulation technique to robot control," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 9866–9872.

[8] J. A. Marvel, R. Bostelman, and J. Falco, "Multi-robot assembly strategies and metrics," *ACM Computing Surveys (CSUR)*, vol. 51, no. 1, pp. 1–32, 2018.

[9] A. Albu-Schäffer, C. Ott, and G. Hirzinger, "A unified passivity-based control framework for position, torque and impedance control of flexible joint robots," *The international journal of robotics research*, vol. 26, no. 1, pp. 23–39, 2007.

[10] T. Inoue, G. De Magistris, A. Munawar, T. Yokoya, and R. Tachibana, "Deep reinforcement learning for high precision assembly tasks," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 819–825.

[11] N. Lin, L. Zhang, Y. Chen, Y. Zhu, R. Chen, P. Wu, and X. Chen, "Reinforcement learning for robotic safe control with force sensing," in *2019 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*. IEEE, 2019, pp. 148–153.

[12] C. Gaz, M. Cognetti, A. Oliva, P. R. Giordano, and A. De Luca, "Dynamic identification of the franka emika panda robot with retrieval of feasible parameters using penalty-based optimization," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4147–4154, 2019.

[13] M. A. Lee, Y. Zhu, K. Srinivasan, P. Shah, S. Savarese, L. Fei-Fei, A. Garg, and J. Bohg, "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks," *arXiv preprint arXiv:1810.10191*, 2018.

[14] J. Luo, E. Solowjow, C. Wen, J. A. Ojea, A. M. Agogino, A. Tamar, and P. Abbeel, "Reinforcement learning on variable impedance controller for high-precision robotic assembly," *arXiv preprint arXiv:1903.01066*, 2019.

[15] G. Schoettler, A. Nair, J. Luo, S. Bahl, J. A. Ojea, E. Solowjow, and S. Levine, "Deep reinforcement learning for industrial insertion tasks with visual inputs and natural rewards," *arXiv preprint arXiv:1906.05841*, 2019.

[16] J. Grundmann, "Design and performance requirements for fuel recycle manipulation systems," in *Performance Evaluation of Programmable Robots and Manipulators: Report of a Workshop Held at Annapolis, Maryland, October 23-25, 1975*, vol. 459. US Department of Commerce, National Bureau of Standards, 1976, p. 147.

[17] E. Helms, R. D. Schraft, and M. Hagele, "rob@ work: Robot assistant in industrial environments," in *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*. IEEE, 2002, pp. 399–404.

[18] M. Hvilshøj, S. Bøgh, O. Madsen, and M. Kristiansen, "The mobile robot little helper: concepts, ideas and working principles," in *2009 IEEE Conference on Emerging Technologies & Factory Automation*. IEEE, 2009, pp. 1–4.

[19] C. Wurll, T. Fritz, Y. Hermann, and D. Hollnaicher, "Production logistics with mobile robots," in *ISR 2018; 50th International Symposium on Robotics*. VDE, 2018, pp. 1–6.

[20] N. Krüger, A. Ude, H. G. Petersen, B. Nemec, L.-P. Ellekilde, T. R. Savarimuthu, J. A. Rytz, K. Fischer, A. G. Buch, D. Kraft *et al.*, "Technologies for the fast set-up of automated assembly processes," *KI-Künstliche Intelligenz*, vol. 28, no. 4, pp. 305–313, 2014.

[21] M. A. Lee, Y. Zhu, K. Srinivasan, P. Shah, S. Savarese, L. Fei-Fei, A. Garg, and J. Bohg, "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8943–8950.

[22] Y. Zheng, X. Zhang, Y. Chen, and Y. Huang, "Peg-in-hole assembly based on hybrid vision/force guidance and dual-arm coordination," in *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2017, pp. 418–423.

[23] Y. Huang, X. Zhang, X. Chen, and J. Ota, "Vision-guided peg-in-hole assembly by baxter robot," *Advances in Mechanical Engineering*, vol. 9, no. 12, p. 1687814017748078, 2017.

[24] R. Li and H. Qiao, "A survey of methods and strategies for high-precision robotic grasping and assembly taskssome new trends," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 6, pp. 2718–2732, 2019.

[25] A. Al-Zabt and T. A. Tutunji, "Robotic arm representation using image-based feedback for deep reinforcement learning," in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*. IEEE, 2019, pp. 168–173.

[26] R. Bogue, "The role of artificial intelligence in robotics," *Industrial Robot: An International Journal*, 2014.

[27] J. Luo, E. Solowjow, C. Wen, J. A. Ojea, and A. M. Agogino, "Deep reinforcement learning for robotic assembly of mixed deformable and rigid objects," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2062–2069.

[28] M. A. Lee, C. Florensa, J. Tremblay, N. Ratliff, A. Garg, F. Ramos, and D. Fox, "Guided uncertainty-aware policy optimization: Combining learning and model-based strategies for sample-efficient policy learning," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 7505–7512.

[29] N. Hogan, "Impedance control: An approach to manipulation: Part itheory," 1985.

[30] J. E. Colgate and N. Hogan, "Robust control of dynamically interacting systems," *International journal of Control*, vol. 48, no. 1, pp. 65–88, 1988.

[31] D. E. Whitney, "Force feedback control of manipulator fine motions," 1977.

[32] E. Matheson, R. Minto, E. G. Zampieri, M. Faccio, and G. Rosati, "Human–robot collaboration in manufacturing applications: A review," *Robotics*, vol. 8, no. 4, p. 100, 2019.

[33] S. Su, X. Zeng, S. Song, M. Lin, H. Dai, W. Yang, and C. Hu, "Positioning accuracy improvement of automated guided vehicles based on a novel magnetic tracking approach," *IEEE Intelligent Transportation Systems Magazine*, 2018.

[34] D. Kragic, H. I. Christensen *et al.*, "Survey on visual servoing for manipulation," *Computational Vision and Active Perception Laboratory, Fiskartorpsv*, vol. 15, p. 2002, 2002.

[35] Q. Wei, C. Yang, W. Fan, and Y. Zhao, "Design of demonstration-driven assembling manipulator," *Applied Sciences*, vol. 8, no. 5, p. 797, 2018.

[36] S. Gottschlich, C. Ramos, and D. Lyons, "Assembly and task planning: A taxonomy," *IEEE Robotics & Automation Magazine*, vol. 1, no. 3, pp. 4–12, 1994.

[37] T. Johannink, S. Bahl, A. Nair, J. Luo, A. Kumar, M. Loskyll, J. A. Ojea, E. Solowjow, and S. Levine, "Residual reinforcement learning for robot control," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 6023–6029.

[38] V. Lippiello, B. Siciliano, and L. Villani, "Position-based visual servoing in industrial multirobot cells using a hybrid camera configuration," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 73–86, 2007.

[39] R. Y. Tsai, R. K. Lenz *et al.*, "A new technique for fully autonomous and efficient 3 d robotics hand/eye calibration," *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 345–358, 1989.

[40] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," *arXiv preprint arXiv:1511.05952*, 2015.

[41] A. Chan, E. A. Croft, and J. J. Little, "Constrained manipulator visual servoing (cmvs): Rapid robot programming in cluttered workspaces," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2011, pp. 2825–2830.

[42] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE transactions on robotics and automation*, vol. 12, no. 5, pp. 651–670, 1996.

[43] H. Park, J. Park, D.-H. Lee, J.-H. Park, and J.-H. Bae, "Compliant peg-in-hole assembly using partial spiral force trajectory with tilted peg posture," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4447–4454, 2020.