Learned impedance policies facilitate robotic assembly in general direction and integration of visual sensor

Eylon Cohen[1], Sher Hazan[1], Ronit Schneor[1], Anath Fischer[2] and Miriam Zacksenhouse[2] *Member, IEEE*

***Abstract*—Impedance control has been demonstrated to facilitate robotic assembly under uncertainties. Position based impedance control, also known as admittance control, involves modifying the reference trajectory in response to force and torque (F/T) measurements, to satisfy the desired dynamical behavior. However, previous work focused on cases when the F/T sensor is co-aligned with the direction of insertion. In those cases, when the mating parts overlap, the direction of the torque vector provides unambiguous information about the relative location of the parts. However, this relationship does not hold when the F/T sensor is not co-aligned with the direction of insertion. Here we propose a method for impedance control independent of the alignment between the F/T sensor and the direction of insertion. The method is based on transforming the F/T measurements to a virtual location that further simplifies the impedance control, and on replacing the noisy friction forces with virtual friction forces. The proposed method facilitates the integration of learning-based visual sensor when a wrist-mounted camera is available. The impedance parameters were learned in simulation using reinforcement learning (RL). The proposed method, with the learned parameters, was evaluated both in stimulation and on a physical robot (UR5e). The experiments demonstrate that the proposed method performs successfully despite uncertainties, and generalizes well to different sizes and even to assembly of flexible parts. Finally, the visual sensor improved performance up to 100% success rate on the physical robot. Thus, the proposed method can facilitate the integration of autonomous robots in industrial assembly.**

***Index Terms*—Compliance control, impedance control, robotic control, robotic assembly, machine learning, reinforcement learning, force and torque sensor.**

I. Introduction

ROBOTS are robust, precise, programmable, and repeatable machines that can be designated for a large variety of tasks [1]. Hence, there has been a growing interest in transferring tasks currently performed by humans to autonomous robots [2]. Indeed, over the last few decades robots have been integrated into a range of industries, especially large industries (25% of Europe's large enterprises use robots [3]).

One of the main challenges for robots is the execution of contact-rich assembly tasks, since they sensitive to uncertainties in the pose of the parts. Nevertheless, assembly tasks are common and necessary in a wide range of industries, including medical devices, heavy industry, and electronics.

A common approach to overcome uncertainties is to control the position of the mating parts precisely, but, in many cases, this is not cost-effective. Hence, assembly tasks are mainly performed by humans (only 2% of Europe enterprises use robots for assembly tasks [3]). Robots that could overcome uncertainties in assembly tasks are expected to increase production rate, lower costs, and significantly reduce the risk of human injury.

Impedance control is a well-known control strategy for tasks involving interaction with the environment [4]. Position based impedance contorl, also known as admittance control, modifies the reference trajectory in response to force and torques (F/T) measurements, according to the desired dynamical behavior. However, the application of impedance control when the F/T sensor is not co-aligned with the direction of insertion is challenging, as explained in Section III-B.

Here we address the challenge of performing assembly tasks despite uncertainties, when the F/T sensor is not co-aligned with the direction of insertion. Furthermore, we address the following challenges: (1) generalization over sizes, (2) generalization to assembly of flexible parts, and (3) integration of visual information to improve performance when a wrist camera is available.

The proposed method relies on transforming the measurements of an actual F/T sensor to a virtual location. That location is selected so the resulting virtual torque can be converted directly to virtual friction forces, which facilitate impedance control. We use position-based impedance control and learn the parameters of the impedance matrices using reinforcement learning (RL).

The work was motivated by the challenge of horizontal assembly of flexible medical tubes of $8.5[mm]$ diameter (MD industries [5]). To avoid the simulation of flexible tubes, training was based on the simulation of assembly of rigid parts of similar sizes. Using the learned parameters, the proposed methods (with and without a visual sensor) were evaluated on a physical robot, UR5e, with both rigid parts and flexible medical tubes (see video).

*The work was supported by the Israel Innovation Authority as part of the ART (Assembly by Robotic Technology) consortium (grant number 74885).

[1] Eylon Cohen, Sher Hazan, and Ronit Schneor are with the Faculty of Mechanical Engineering, Technion, Israel. ela3135@gmail.com, sherhazan1115@gmail.com, schneor@me.technion.ac.il

[2]Anath Fischer and Miriam Zacksenhouse are with the Faculty of Mechanical Engineering, Technion, Israel and the Technion Autonomous Systems Program. meranath@technion.ac.il, mermz@technion.ac.il

II. Related Work

Impedance control endows the end-effector (EEF) with the desired impedance, e.g., stiffness, damping and inertia, to achieve the desired trade-off between position deviations and force deviations during interactions with the environment. However, to date, the applications of impedance control to assembly tasks focused mainly on cases when the F/T sensor is co-aligned with the direction of insertion ([6], [7], [8], [9]). In those cases, when the mating parts overlap, the direction of the resulting torque depends on the direction of the overlap [10], and thus can be used to determine the required correction.

However, when the F/T is not co-aligned with the direction of insertion, the direction of the torque vector does not depend on the relative location of the parts, as further explained in Section III-B. Nevertheless, this case is common in daily life and industrial applications due to spatial arrangements and space limitations.

In robotic applications impedance control is usually implemented in software rather than in hardware. The standard dynamic based impedance control introduced by Hogan [11] relies on an accurate dynamic model of the robot [12], [13] and thus may hamper sim2real. Instead, we implemented position based impedance control, also known as admittance control, which modifies the reference trajectory in response to F/T measurements [13], [14].

Following our previous work on vertical assembly using impedance policies [8], [9], we learn a fixed set of impedance parameters using RL. However, while our previous work demonstrated the importance of using non-diagonal and especially asymmetric impedance matrices, here we use diagonal matrices. The diagonal matrices are sufficient in the proposed method since we introduce virtual friction forces, which account for the effect of torques, as further explained in Section III-C. Thus, the proposed method greatly reduces the number of free parameters that have to be learned by RL.

III. Reinforcement Learning of Impedance Policies

A. Position Based Impedance Control

Position based impedance control modifies the desired trajectory X^w_{dw}, specified by the position vector ${p^w_{dw}}^{3\times1}$ and rotation matrix ${R^w_{dw}}^{3\times3}$ with respect to the world frame, to a compliant trajectory X^w_{cw}, specified by ${p^w_{cw}}^{3\times1}$ and ${R^w_{cw}}^{3\times3}$. Using the angle/axis representation of the relative orientation of the compliant and desired trajectories [15], and assuming small relative angle, the 6 degrees of freedom impedance model can be expressed as:

$$M^{6\times6}\ddot{x}^{d\,6\times1}_{cd} + C^{6\times6}\dot{x}^{d\,6\times1}_{cd} + K^{6\times6}x^{d\,6\times1}_{cd} = F^{d\,6\times1} \quad (1)$$

where: $x^d_{cd} = \left[{p^d_{cd}}^T, {o^d_{cd}}^T\right]^T$ is the relative position p^d_{cd} and the small relative orientation o^d_{cd} of the compliant trajectory with respect to the desired trajectory (expressed in the frame attached to the desired trajectory), $\dot{x}^d_{cd} = \left[{p^d_{cd}}^T, {\omega^d_{cd}}^T\right]^T$ is the translation and angular velocities between the trajectories, expressed in the frame attached to the desired trajectory, $M, C, K \in R^{6\times6}$ are the desired inertia, damping and stiffness of the impedance model, respectively, and ${F^d}^{6\times1}$ is the force and torque vector.

Eq. (1) can be expressed in state space representation:

$$\dot{X}^d_{cd} = AX^d_{cd} + BU^d \quad (2)$$

where:

$${X^d_{cd}}^{12\times1} = \begin{bmatrix} x^d_{cd} \\ \dot{x}^d_{cd} \end{bmatrix} = \begin{bmatrix} p^d_{cd} \\ o^d_{cd} \\ \dot{p}^d_{cd} \\ \omega^d_{cd} \end{bmatrix} \;\&\; {\dot{X}^d_{cd}}^{12\times1} = \begin{bmatrix} \dot{p}^d_{cd} \\ \omega^d_{cd} \\ \ddot{p}^d_{cd} \\ \dot{\omega}^d_{cd} \end{bmatrix}, \quad (3)$$

$$A^{12\times12} = \begin{bmatrix} 0^{6\times6} & I^{6\times6} \\ -{M^{-1}}^{6\times6}K^{6\times6} & -{M^{-1}}^{6\times6}C^{6\times6} \end{bmatrix}, \quad (4)$$

$$B^{12\times6} = \begin{bmatrix} 0^{6\times6} \\ {M^{-1}}^{6\times6} \end{bmatrix} \;\&\; U^{d\,6\times1} = F^{d\,6\times1} \quad (5)$$

Finally, since realistic control executes in discrete time steps of ΔT, a discrete state space realization is implemented:

$$X^d_{cd}(k+1) = A_d X^d_{cd}(k) + B_d U^d(k) \quad (6)$$

where:

$$A_d = e^{A\Delta T} \;\&\; B_d = A^{-1}(A_d - I)B \quad (7)$$

To assure stability we verified that the eigenvalues of A_d are inside the unit circle.

The desired trajectory X^w_{dw} was designed as a minimum jerk trajectory to the desired final position and orientation. The compliant trajectory in the world frame, X^w_{cw}, was computed from the desired trajectory and the deviation X^d_{cd}.

B. Virtual Sensor

When the peg partially overlaps the hole, the reaction forces are skewed to one side of the peg. If the F/T sensor is co-aligned with the direction of the insertion, as in Figure 1(a), the direction of the resulting torque depends on the relative position of the peg and hole [10]. In that case, the non-diagonal terms of the stiffness matrix can help determining the necessary correction movements [9]. However, when the F/T sensor is not co-aligned with the direction of insertion, as in Figure 1(b), the direction of the torque may not vary with the relative location of the peg and hole.

To overcome this problem, this study suggests to transform the measured F/T to a virtual location, thus generating a virtual sensor (Figure 1(c)). In general, the virtual sensor can be located in any place, but, as will be justified latter, we suggest to position it at the center of the edge of the peg, facing the hole, as indicated in Figure 1(c).

In order to derive the transformation between the measured and virtual F/T sensors, we first derive the relationship between the measured and external F/T. For the static and gravity-free case, the force, F_S and torque T_S, measured by

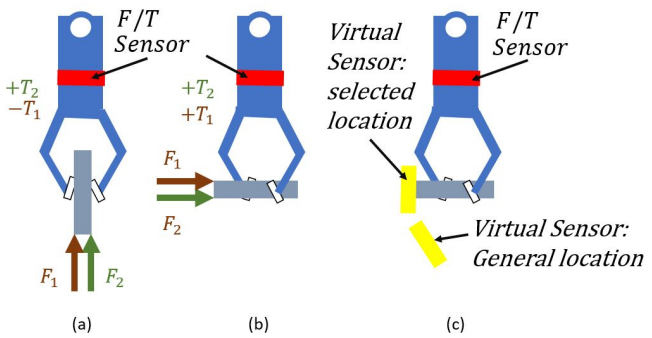


Fig. 1: Effect of the alignment of the F/T sensor with the direction of insertion in a 2-dimensional case ((a) and (b)) and possible locations of the virtual F/T sensor (c). Co-aligned sensor (a): the sign of the torque depends on the side of the peg on which the reaction force is applied. Non co-aligned sensor (b): the sign of the torque is the same independent of where the reaction force is applied. Virtual sensor can be located in different places, including the selected place at the center of the the edge of the peg facing the hole (c).

the F/T sensor located at P_S are related to the external force F_{ext} that acts at P_{ext} and to the external torque T_{ext} by:

$$\begin{bmatrix} F_S \\ T_S \end{bmatrix} = \begin{bmatrix} F_{ext} \\ (P_{ext} - P_S) \times F_{ext} + T_{ext} \end{bmatrix} \tag{8}$$

Similarly, a virtual sensor in some virtual position P_V would have measured the force F_V and torque T_V given by:

$$\begin{bmatrix} F_V \\ T_V \end{bmatrix} = \begin{bmatrix} F_{ext} \\ (P_{ext} - P_V) \times F_{ext} + T_{ext} \end{bmatrix} \tag{9}$$

Given that:

$$P_{ext} - P_V = (P_S - P_V) + (P_{ext} - P_S), \tag{10}$$

Eq. (9) can be expressed as:

$$\begin{bmatrix} F_V \\ T_V \end{bmatrix} = \begin{bmatrix} F_{ext} \\ ((P_S - P_V) + (P_{ext} - P_S)) \times F_{ext} + T_{ext} \end{bmatrix} \tag{11}$$

Finally, inserting Eq. (8) in Eq. (11) results in:

$$\begin{bmatrix} F_V \\ T_V \end{bmatrix} = \begin{bmatrix} F_S \\ (P_S - P_V) \times F_S + T_S \end{bmatrix} \tag{12}$$

Eq. (12) implies that, for the static and gravity-free case, the F/T measured by the actual sensor can be transformed to the F/T measured at any virtual sensor given the relative position of the actual and virtual sensors. For the general case where the body's dynamic and gravity are not negligible, it can be shown [16] that those elements can be compensated in real-time.

For reasons detailed in the next section, we placed the virtual sensor at the center of the bottom surface of the peg. The position of the virtual sensor was determined in a series of calibration experiments, and possible calibration errors were considered in the training process as detailed in Section III-E.

C. Virtual Friction Forces

We suggest to locate the virtual sensor such that the friction forces, which tend to by noisy, do not affect the virtual torque. Hence, we locate the virtual sensor at the center of the edge of the peg facing the hole and align its z-axis with the direction of insertion. In that case, the vector $(P_S - P_V)$ lies on the same plane as the friction Forces, and thus the x and y components of the virtual torque, T_{V_x} and T_{V_y}, are not affected by them.

When the peg contacts the surface of the hole (see Figure 2), the 2-dimensional vector R from the virtual sensor to the location of the equivalent reaction force F_{V_z}, is given by:

$$R = \begin{bmatrix} -\frac{T_{V_y}}{F_{V_z}} & \frac{T_{V_x}}{F_{V_z}} \end{bmatrix}^T \tag{13}$$

As is evident from Figure 2, under contact, the vector R should be along the line connecting the centers of the hole and peg. Hence, we suggest to replace the measured friction forces, which tend to be noisy, with an equivalent unit virtual friction force $\hat{F} = -\frac{R}{|R|}$. This replacement is expected to facilitate the ability of the impedance controller to modify the reference trajectory in the direction of the hole center.

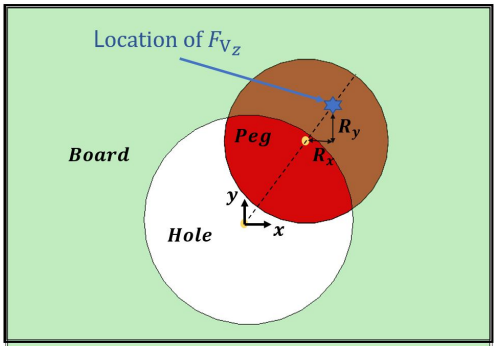


Fig. 2: Illustration of the contact surface between the peg and the hole. The blue star represents the location of the equivalent reaction force F_{V_z} acting on the peg and R represents its position relative to the center of the peg. Ideally, it is assumed that R is on the line connecting the centers of the peg and hole.

Thus, the right hand side of the impedance model (Eq. (1)) is replaced by $Q^{6\times1} = [\hat{F}, F_{V_z}, T_{V_x}, T_{V_y}, T_{V_z}]^T$, resulting in the control method illustrated in Figure 3.

D. Action space

In the RL framework, the agent interacts with its environment by performing actions. As in our previous work [9], the action in our method determines the parameters of the impedance matrices K, C, M, which in turn affects the trajectory correction. In [9] it was demonstrated that it is sufficient to learn space-invariant impedance parameters. Thus, here we focus on learning a fixed set of parameters.

We hypothesized that given the virtual friction forces, it would be sufficient to use diagonal impedance matrices rather than the asymmetric matrices used in [9]. As explained

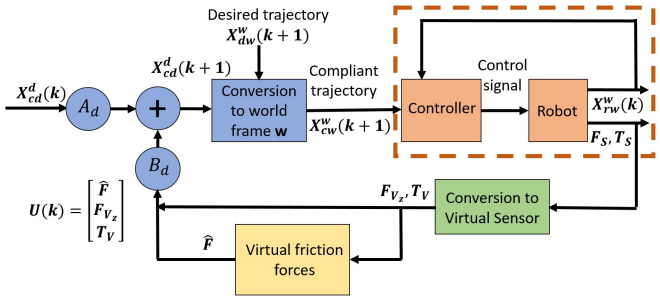


Fig. 3: High-level control scheme: position based impedance control modifies the desired trajectory X^w_{dw}, specified in the world frame, w, to a compliant trajectory X^w_{cw}, such that its deviation from the desired trajectory, X^d_{cd}, satisfies the impedance model (captured by A_d and B_d) in response to the virtual F/T measurements (F_V and T_V) and virtual friction forces ($\hat{F}$). The compliant trajectory is the input to the low-level controller marked in a dashed orange frame, that also receives the robot state X^w_{rw}. The virtual F/T measurements are computed from the actual F/T measurements (F_S and T_S). The subscripts d, c, w and r refer to the desired, compliant, world and robot, respectively

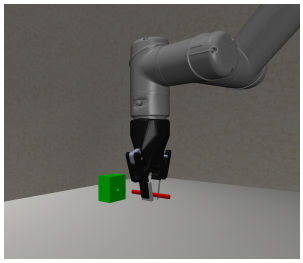


Fig. 4: Simulation of the horizontal insertion task showing a close up of the gripper, the peg and the board with a hole.

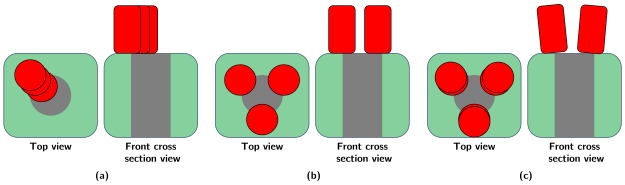


Fig. 5: Types of possible errors that were deliberately introduced in the simulation to train a policy that is robust to uncertainties. Top view (right) and front cross-section view (left) of radial errors (a), angular error (b), and peg orientation errors (c).

and demonstrated in [9], asymmetric impedance matrices are important in order to modify the x and y position in response to T_y and T_x, respectively. However, the virtual friction force $\hat{F}$, which is a unit vector in the direction of R given by Eq. (13), reflects already the effect of the torque. Thus, we restricted the impedance matrices to be diagonal.

Finally, when the parts are axisymmetric (i.e., have round cross-section), the structure of each impedance matrix can be further restricted to:

$$*^{6\times 6} = diag(\begin{bmatrix} *_{xy} & *_{xy} & *_z & *_{\theta_x\theta_y} & *_{\theta_x\theta_y} & *_{\theta_z} \end{bmatrix}) \quad (14)$$

where $*$ stands for K, C or M. Thus, each matrix is defined by 4 parameters, for a total of only 12 free parameters.

E. Learning the impedance policy in Simulation

The simulation environment was implemented using robosuite [17], a well-known, open-source, physical simulation platform powered by the MuJoCo [18] physical engine. The impedance policy, which determines the parameters of the impedance matrices, was trained to perform the nominal task specified below, using Proximal Policy Optimization (PPO) algorithm [19], implemented in Stable baseline 3 [20].

Nominal task: The policy was trained to perform horizontal insertion of a circular peg with diameter $D_{peg} = 8[mm]$ into a round hole with diameter $D_{hole} = 10[mm]$ located on a board mounted vertically on a table, as depicted in Figure 4. The peg was held by the gripper so it was horizontal to the table. For simplicity, the location of the board was fixed for all episodes (as explained above, and tested on the physical robot, this does not hamper the generalization to other locations).

In facilitate robustness to uncertainties in the location of the hole and orientation of the peg, and to calibration errors (see end of Section III-B), the following errors were deliberately introduced during training:

- Radial error R_{err} and angular error A_{err} between the actual and presumed locations of the hole (Figure 5 (a) and (b), respectively) were uniformly distributed in the ranges $[1.5, 3.5]$[mm] and $[0, 2\pi]$[rad], respectively
- Peg orientation error O_{err} (Figure 5 (c)) was uniformly distributed in the range $[-5°, 5°]$.
- Axial and radial calibration errors in the position of the virtual sensor, A_{cal} and R_{cal}, respectively, were uniformly distributed in the ranges $[-5, 5]$[mm] and $[0, 2]$[mm,] respectively.

Reward function: The reward was composed of two parts: (1) A step-by-step penalty that depended on the relative location and orientation of the peg and hole, and (2) A final bonus, $R_B = 10^5$, that was given for successful insertions

F. Visual Friction Forces

A wrist camera can be used to improve performance by estimating the direction of the virtual friction forces visually. For that purpose, we defined the angle α, between the vector from the center of the hole to the center of the peg and the x-axis. The angle α was estimated from the visual sensor and used to define the visual friction force $\hat{F} = -\begin{bmatrix} cos(\alpha) & sin(\alpha) \end{bmatrix}^T$, which was used instead of the virtual friction force.

The relative angle α was estimated using a classification-based approach. First, the spatial location was discretized into several angular ranges according to the desired resolution (here we used 16 ranges). Second, a ResNet-18 [21] network was trained to classify the angular range. Finally, the output

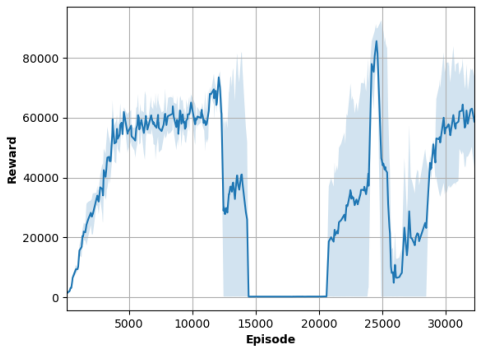


Fig. 6: Average reward during a training session of horizontal insertion with random errors. Blue background denotes the standard deviation envelope.

of the network was interpreted as the likelihood of each range, and α was estimated using linear interpolation.

IV. Results

A. Simulation Results

Figure 6 depicts the average episode reward during a typical training session described in Section III-E. Recall that the maximum possible reward is $R_B = 10^5$. While the reward increased initially, it dropped sharply before increasing again to its maximum level. This might be attributed to the drastic effect of a slight change in the policy on the impedance.

The performance of the proposed method, with the parameters that resulted in the highest reward during training, was evaluated in simulation on both the nominal task and on two new tasks. The new tasks evaluated generalization over size. Each task was evaluated with and without calibration errors (R_{cal} and A_{cal}) for 200 trials, and the resulting success rates and their confidence intervals (CI) are summarized in Table I.

Table I indicates that the estimated success rates remained above 91% for all tasks, thus demonstrating robustness to uncertainties and generalization over size. Interestingly, if the location of the virtual sensor is accurate ($R_{cal} = A_{cal} = 0$), the estimated success rate of inserting a 4.5[mm] peg into a 6[mm] hole was 100%, despite uncertainties in the radial position of the hole (uniformly distributed in the range of $1-2.5$[mm]). The considered calibration errors and uncertainties in pole orientation degraded performance by $3.5-7.5\%$.

B. Simulation to Real Transfer (Sim2Real)

The proposed method for horizontal insertion, with the parameters learned in the simulation, was implemented on a real robot, UR5e, equipped with an OnRobot HEX-E F/T sensor. Performance on the real robot was evaluated for the same sizes and uncertainties in hole position as in the simulations. Uncertainties in the position of the virtual sensor and the orientation of the peg resulted naturally from the estimation of the location of the virtual sensor and from

D_{peg} [mm]	D_{hole} [mm]	R_{err} [mm]	A_{cal}/R_{cal} [mm]	O_{err} [°]	Success rate [%]	95% CI [%]
8	**10**	**[1.5,3.5]**	0/0	0	97	[93.6,98.6]
			[-5,5]/[0,2]	**[-5,5]**	**93.5**	**[89.2,96.2]**
16	20	[2.5,8]	0/0	0	94	[89.8,96.5]
		[2.5,5]	[-5,5]/[0,2]	[-5,5]	91	[86.2,94.2]
4.5	6	[1,2.5]	0/0	0	100	[98.1,100]
			[-2,2]/[0,1]	[-5,5]	92.5	[88,95.4]

TABLE I: Simulation performance of the trained policy. Success rates and confidence intervals (CI) were estimated from 200 trials of the indicated horizontal insertion task. D_{peg} and D_{hole} are the diameters of the peg and hole, respectively. Errors were uniformly distributed in the indicated ranges in the radial position of the hole R_{err}, peg orientation O_{err} and axial and radial position of the virtual sensor A_{cal} and R_{cal}, respectively. Performance of the nominal task is marked in bold (second row).

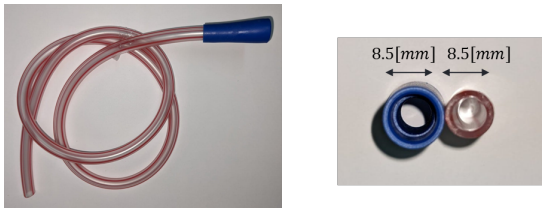


Fig. 7: Flexible medical tubes that were inserted horizontally during experiments on the real robot (MDC industries).

grasping the peg. The experimental results are summarized in the first 3 rows of Table II. The transfer to the real robot degraded performance moderately (by $10.5\% - 11\%$) for the nominal and large pegs (first two rows), and drastically (by 41.5%) for the small peg (third row). This degradation can be attributed to measurement noise in the real sensor and possibly larger calibration errors (in the position of the virtual sensor).

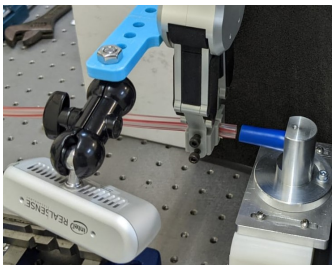
We also tested the generalization of the method to assembly of flexible parts, and in particular, to the assembly of flexible medical tubes (MDC industries [5]). The diameter of both mating tubes was $8.5[mm]$, as depicted in Figure 7. As reported in the last row of Table II (marked in blue), the success rate in assembling the flexible tubes with the real robot was 80%, despite uncertainties in the relative position and the flexibility of the parts.

C. Optional Visual Sensor

A REALSENSE D435 camera was mounted on the gripper of the UR5e, as shown in Figure 8. Separate sets of training images were collected with rigid parts and flexible tubes. The gripped part was placed at uniformly distributed angles and radii around the fixed part: radius of up to $(D_{hole} + D_{peg})/2$ for the rigid parts, and $1.5-6$[mm] for the tubes. Separate networks were trained for the rigid parts and flexible tubes. The trained networks predicted α with mean absolute error (MAE) of $5.98°$ for the rigid parts and $7.33°$ for the flexible tubes (root mean square, RMS, of $8.5°$ and $11.5°$, respectively).

D_{peg} [mm]	D_{hole} [mm]	R_{err} [mm]	Success rate [%]	95% CI [%]
8	**10**	**[1.5, 3.5]**	**83**	**[74.5, 89.1]**
16	20	[2.5, 8]	80	[71.1, 86.7]
4.5	6	[1, 2.5]	51	[41.3, 60.6]
8.5	8.5	[1.5, 2.5]	80	[71.1, 86.7]

TABLE II: Experimental performance of the trained policy, with virtual friction forces, on a physical robot UR5e equipped with OnRobot F/T sensor. Success rates and confidence intervals (CI) were estimated from 100 trials of the specified horizontal insertion task. See Table I for other symbols. Performance of the nominal task is marked in bold. Performance with flexible tubes is marked in blue (last row).



(a)

(b)

Fig. 8: Camera mounted on the end effector during insertion of medical tubes (a) and rigid cylindrical peg (b).

The performance of the impedance controller with both the virtual and visual sensors, was evaluated on the physical robot, as summarized in Table III. Perfect success rates of 100% were achieved for the nominal task on which training was conducted, and even for the larger peg and hole (16[mm] peg into 20[mm] hole), over a large range of position uncertainties (larger than the range used in training the impedance policy). The success rate for inserting the 8.5[mm] medical tube was 98%, with radial uncertainties of 2[mm] to 6[mm]. Comparing these results to the ones to without the visual sensor (Table II), it is evident that integrating the visual sensor increased the success rate over a broader range of uncertainties.

V. Conclusion

The goal of this research was to develop a robust method for peg-in-hole insertion tasks, regardless of whether the direction of insertion is co-aligned with the F/T sensor. In particular, the work focused on horizontal insertion with a vertically aligned gripper. Given the importance of impedance control for contact-rich tasks, and previous success in our laboratory, we developed a method based on impedance control. To handle cases where the insertion axis is not co-aligned with the axis of the F/T sensor, we introduced a virtual sensor located at the center of the edge of the peg, facing the hole. This facilitated the computation of virtual friction forces from the measured torques, instead of the inherently noisy friction forces. The virtual friction force directly relates the torques to linear path corrections and so even diagonal impedance matrices with only 12 free parameters worked well.

D_{peg} [mm]	D_{hole} [mm]	R_{err} [mm]	Success rate [%]	95% CI [%]
8	**10**	**[1.5, 10]**	**100**	**[96.3, 100]**
16	20	[4, 14]	100	[96.3, 100]
4.5	6	[1, 3]	90	[82.6, 94.5]
8.5	8.5	[2, 6]	98	[93, 99.4]

TABLE III: Experimental performance of the trained policy, with visual friction forces, on a physical robot UR5e equipped with OnRobot F/T sensor and REALSENSE D435 camera. Success rates and confidence intervals (CI) were estimated from 100 trials of the specified horizontal task. See Table I for other symbols. Performance of the nominal task, with larger uncertainties, is marked in bold. Performance with the flexible tubes is marked in blue (last row).

RL was used to learn the impedance parameters for a nominal horizontal insertion of rigid parts, despite uncertainties in the position. Simulation results indicate good success rate (93%) for the nominal task, and good generalization across size, especially to larger sizes. The transition from simulation to the physical robot degraded performance by about 10% for the nominal task.

The method generalized very well to flexible pegs, and achieved a good success rate (80%) for inserting $8.5[mm]$ medical tubes. The generalization to flexible tubes is especially important since flexible parts are difficult to simulate well. Our method demonstrate that it is enough to learn the impedance policy in simulations of rigid parts of similar sizes - and the learned policy perform well on flexible parts. Nevertheless, future work may evaluate the effect of retraining on the real robot.

We have shown that the method can be easily enhanced using a visual sensor when available by computing visual, rather than virtual, friction forces. Using the visual sensor, success rates on the real robot increased to 100% for the nominal task, and for larger sizes, and to 98% for assembly of flexible medical tubes, despite larger ranges of uncertainties.

In conclusion, the proposed method facilitates robotic assembly in general directions, independent of the axis of the F/T sensor, and achieves good success rates despite position uncertainties. It generalizes well to larger sizes, and most importantly, to assembly of flexible parts like medical tubes. Finally, the proposed method can take advantage of a wrist-mounted visual sensor and achieve close to 100% success rate. Thus, the proposed method can facilitate the integration of autonomous robots in industrial assembly.

REFERENCES

[1] T. Turja and A. Oksanen, "Robot acceptance at work: a multilevel analysis based on 27 eu countries," *International Journal of Social Robotics*, vol. 11, no. 4, pp. 679–689, 2019.

[2] G. Michalos, S. Makris, J. Spiliotopoulos, I. Misios, P. Tsarouchi, and G. Chryssolouris, "Robo-partner: Seamless human-robot cooperation for intelligent, flexible and safe operations in the assembly factories of the future," *Procedia CIRP*, vol. 23, pp. 71–76, 2014.

[3] P. E. News, "25% of large enterprises in the eu use robots," https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20190121-1, 2019.

[4] N. Hogan, "Impedance control: An approach to manipulation: Part ii—implementation," 1985.

[5] Mdc, "Mdc industries," http://www.mdcindustries.com, 2022.

[6] S. Chan and H. Liaw, "Generalized impedance control of robot for assembly tasks requiring compliant manipulation," *IEEE Transactions on Industrial Electronics*, vol. 43, no. 4, pp. 453–461, 1996.

[7] J. F. Broenink and M. L. Tiernego, "Peg-in-hole assembly using impedance control with a 6 dof robot," in *Proceedings of the 8th European Simulation Symposium*. Citeseer, 1996, pp. 504–508.

[8] O. Spector and M. Zacksenhouse, "Deep reinforcement learning for contact-rich skills using compliant movement primitives," *arXiv preprint arXiv:2008.13223*, 2020.

[9] S. Kozlovsky, E. Newman, and M. Zacksenhouse, "Reinforcement learning of impedance policies for peg-in-hole tasks: Role of asymmetric matrices," *IEEE Robotics and Automation Letters*, 2022.

[10] T. Kusakabe, S. Sakaino, and T. Tsuji, "Design of non-diagonal stiffness matrix for assembly task," *arXiv preprint arXiv:2210.16594*, 2022.

[11] N. Hogan, "Impedance Control Part1-3," *Transaction of the ASME, Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. March 1985, pp. 1–24, 1985.

[12] A. Lu, Z.; Goldenberg, "Implementation of Robust Impedance and Force Control," pp. 145–163, 1992.

[13] T. Valency and M. Zacksenhouse, "Accuracy/robustness dilemma in impedance control," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 125, no. 3, pp. 310–319, 2003.

[14] M. Schumacher, J. Wojtusch, P. Beckerle, and O. von Stryk, "An introductory review of active compliant control," *Robotics and Autonomous Systems*, vol. 119, no. April 2020, pp. 185–200, 2019.

[15] F. Caccavale, C. Natale, B. Siciliano, and L. Villani, "Six-dof impedance control based on angle/axis representations," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 2, pp. 289–300, 1999.

[16] J. Duan, Z. Liu, Y. Bin, K. Cui, and Z. Dai, "Payload identification and gravity/inertial compensation for six-dimensional force/torque sensor with a fast and robust trajectory design approach," *Sensors*, vol. 22, no. 2, p. 439, 2022.

[17] Y. Zhu, J. Wong, A. Mandlekar, and R. Martín-Martín, "robosuite: A modular simulation framework and benchmark for robot learning," in *arXiv preprint arXiv:2009.12293*, 2020.

[18] E. Todorov, T. Erez, and Y. Tassa, "Mujoco: A physics engine for model-based control," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 5026–5033.

[19] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[20] A. Raffin, A. Hill, M. Ernestus, A. Gleave, A. Kanervisto, and N. Dormann, "Stable baselines3," 2019.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.