Trajectory Generation Using Dual-Robot Haptic Interface for Reinforcement Learning from Demonstration

Daniel Frau-Alfaro[1](✉), Santiago T. Puente[1], and Ignacio de Loyola Páez-Ubieta[1,2]

[1] AUtomatics, RObotics, and Artificial Vision Lab, University Institute for Computer Research, University of Alicante, San Vicente Alicante, Spain
{daniel.frau,santiago.puente,ignacio.paez}@ua.es

[2] University Institute for Engineering Research, University Miguel Hernández, Elche, Spain
ipaez@umh.es

Abstract. In learning robotics, techniques such as Learning from Demonstrations (LfD) and Reinforcement Learning (RL) have become widely popular among developers. However, this approximations can result in inefficient strategies when it comes to train more than one agent interacting in the same space with several objects and unknown obstacles. To solve this problematic, Reinforcement Learning from Demonstration (RLfD) allows the agent to learn and evaluate its performance from a set of demonstrations provided by a human expert while generalising from them using RL training. In dual-robot applications this approach is suitable for training agents that perform collaborative tasks. For this reason, a dual-robot haptic interface has been designed in order to produce dual manipulation trajectories to feed a RLfD agent. Haptics allows to perform high quality demonstrations following an impedance control approach. Trajectories obtained will be used as positive demonstrations so the training environment can generate automatic ones. As a result, this dual-robot haptic interface will provide a few trajectory demonstrations on dual manipulation in order to train agents using RL strategies. The aim of this research is to generate trajectories with this dual-robot haptic interface to train one or more agents following RLfD paradigms. Results show that trajectories performed with this interface present less error and deviation than others performed with a non-haptic interface, increasing the quality of the training data.

Keywords: dual-robot · haptic interface · demonstrations · reinforcement learning from demonstration

1 Introduction

Dual manipulation is a discipline that needs precise programming and a highly accurate environment definition. Not only the static objects on the manipulation

L. Marques et al. (Eds.): ROBOT 2023, LNNS 976, pp. 444–455, 2024.
https://doi.org/10.1007/978-3-031-58676-7_36

zone must be properly modeled, but both robots must know in which position are as well as how to manipulate an object. For this reason, traditional control architectures may not result appropriate for this kind of scenarios.

A possible approach is Learning from Demonstration (LfD), a technique that consists in teaching an agent of any kind to perform a specific task or movement based on demonstrations on how to do it. This learning strategy is usually used to teach the robot how to do complex tasks first performed by an expert [26]. This approximation makes it easier for developers because it does not require any kind of high specialization and knowledge in programming or robotics [1,19].

However, in manipulation tasks LfD needs a large dataset of trajectories performed by an expert, so one of the main problems of these methods is data generation. This paradigm needs a huge amount of demonstrations in order to perform properly in dynamic environments with obstacles and other agents, so obtaining all the data manually may take a very long time and can be sub-optimal due to human error.

On the other hand, Reinforcement Learning (RL) is a machine training method that consists in rewarding certain behaviour in order to accomplish certain goal, so an agent can learn from experience a set of actions and try to generalise from them. This approach offers a high grade of flexibility because an agent can learn how to develop a task in real environments without the need of a person to indicate which action to do next or determine all possible rules the machine must follow [11,14].

In robotics, this type of training needs to explore a vast space of possibilities and configurations in order to achieve certain grade of behaviour accomplishment. As a consequence, it takes long time and resources to generate all possible scenarios and to explore the horizon of possibilities so that the system can converge to stability, and may not reach that point. Also, transitions between simulation and real environments may be troublesome due to differences between both scenarios [29].

In order to solve this problem, information obtained from demonstrations about object manipulation with multiple robots can be used to generate more trajectories automatically. This way, a RL training can start with an approximation on how to perform the manipulation, reducing the training time exponentially and making it easier to converge to a desired behaviour. Demonstrations from the expert serve as a guide to the learning agent with only few of them. In addition, the RL can generalise to a more optimal behaviour using those trajectories as a reference. This technique is called Reinforcement Learning from Demonstration (RLfD) [18]. This kind of strategies are capable of correcting sub-optimal demonstrations made by the expert, resulting in better results than a human operator can obtain.

The aim of this work is to build a dual-robot haptic interface using an impedance control approach [16] in order to train a RLfD algorithm so it can perform dual manipulation tasks such as grasping or moving without colliding with each other. This way, the system is able to generate trajectories for training RLfD agents in simulation and real environments with two Phantom Omni Bundle devices as masters together with UR5e robotic manipulators as slaves. The proposed system moves away from strategies that use visual or kinestesic data, which need of extra processing to work properly. These trajectories, serving as demonstrations, will be used as a reference to generate more examples automatically for RLfD training. In addition, this interface is able to switch from simulation environment to a real one without need of changing the system, so both virtual and real trajectories can be performed and stored. The main contribution of this paper would be the generation of trajectories with a dual-haptic robot interface to train a RLfD agent.

This paper is organised as follows: Sect. 2 mentions several works related to the topic of this paper, Sect. 3 presents the designed system while Sect. 4 shows some result and discussion of the performance of the system, Sect. 5 discuss the system advantages and limitations. Finally, Sect. 6 deals with future planned work to use the proposed system.

2 Related Work

Learning methods involving RL and LfD have been widely studied among developers, producing a large amount of literature. Teleoperation and haptics presents a similar situation. On this section, related work is presented to contextualizing the topic of this research.

2.1 Collaborative Robots and Teleoperation

In the field of robotic manipulation, some tasks are meant to be performed by two or more robots. Those are collaborative tasks that may involve more than one agent. In these cases, robots must learn how to work together in order to accomplish the same goal. AI strategies allow researchers to teach multiple robots how to behave in a collaborative environment with humans, e.g. the research [6] uses programming from demonstration techniques to indicate a robot how to grasp elements with visual information. In addition, some works such as [27] propose a dual-robot collaborative system capable of manipulating objects synchronously. Other works and researches focus on applying haptic methods with assisted guidance when operating a robot at the same time, e.g. [21]. Moreover, several researches use direct control architectures to send commands to one or more robots using arm tracking and shared control in [13] or applying visual detection to determine the position of both hands in order to send commands to a pair of robots in [7].

Despite the works previously mentioned, studies have shown that teaching the robot using data from teleoperation movements may improve the quality of the training and later performance. It all comes to ensure smoothness and precision within the movements, so acquiring trajectory samples with the robot may produce high quality information [22] and it guarantees the operator's security during data acquisition [8]. In order to obtain the trajectories generated by a human, a teleoperation approach is needed, which comes with a huge variety of possible solutions. Haptic has proven to be a suitable one for controlling robotic arms, as well as for generating training samples destined to perform LfD agents training. The force feedback provided by such architecture allows the operator to execute smooth and precise movements, resulting on higher quality data for training. In works such as [24] researchers present new methods on how to estimate the geometry of an object relying on haptic information with a dual-robotic system. Other works propose a dual teleoperation system to perform surface activities on planar objects using a multi-modal architecture involving Motion Capture (MoCap) and haptic information, e.g. [9]. Also, in [4] researchers study how to improve manipulation of an object using a teleoperation system controlled by wearable masters along with haptic feedback. Studies show that operators become more confident with this type of systems.

2.2 Learning AI Methods

As for the automatic learning methods, several papers have been published on LfD strategies. On them, different researchers present a wide variety of solutions involving this type of algorithms. Some works apply this methods in order to train an agent so it can send movement commands to a robot as in [12], where LfD is used to train a model that detects head movement of a person to control a robot through vision. In addition, other papers like [25] proposes a system that learns movements from electromyography signals to perform complex human-like actions. Also, it is important to consider that most of LfD agents are trained with visual demonstrations of human movements using MoCap, e.g. in [10], where a dual robotic arm system learns how to perform assembly tasks, taught by position and orientation data.

On RL, works such as [15] study obstacle avoidance using an anthropomorphic arm and RL techniques. The agent learns how to perform obstacle avoidance thanks to a combination of Deep Deterministic Policy Gradient (DDPG) and Hindsight Experience Replay (HER), working with two learning agents at the same time. One of them is trained to avoid obstacles while the other ensures that the robot reaches the goal point. Then, on [28] researchers propose a trajectory-planning method using Deep Reinforcement Learning (DRL) and joint angle information along with Cartesian coordinates of a robot. A comparison with Bidirectional Rapidly-Exploring Random Tree (Bi-RRT) algorithm is provided, improving its performance in simulation environments. Other works like [20] focus on solving simulation to real gap in RL applications. This research relies on visual mapping from a real environment to the simulation training. Results show a highly accurate performance on simulation and real environments.

Some works such as [17] point out that using demonstrations as part of a RL training would overcome exploration problems and reduce the time taken to obtain training data. Also, some works use Hierarchical Deep Q-Networks (H-DQN) and data augmentation to perform LfD training with higher results than other architectures e.g. Deep Q-Networks (DQN) [3]. As a consequence, positive demonstrations generated by an expert become the starting point for a RL training, reducing the training time and exploration. In addition, in [23] a combination of a RL and a LfD system is developed to teach a robotic arm certain motor skills related with non-prehensile manipulation. This research uses kinesthetic trajectories to start RL training.

3 System Description

The system is built upon a dual-robot haptic interface that allows the operator not only to control a pair of robots remotely and manipulate objects with them, but to generate and store trajectory samples that will be later used to perform a RLfD training using DQN architectures.

As for the haptic control loop, a Cartesian Position-Position method is used. This architecture uses global position references sent to the robot, with positions from the haptic device being up-scaled to match the robot's workspace dimensions. This data is computed internally, providing torque references to each motor using individual joint controllers. However, the masters and slaves present different kinematics, so joint references are not suitable for direct control. For this reason, Cartesian references between master and slave's Tool Central Point (TCP) are used as control commands. Then, using the Denavit - Hartenberg (DH) model [5] of the robot, the Cartesian commands are transformed to joint references that are later applied to each robot joint. In order to avoid disturbances due to instabilities in the operator movement, a median filter is applied to this last control step, before joint commands are sent to the robot.

Simultaneously, the impedance feedback force loop takes place and force information is sent to the operator. Due to kinematics differences previously mentioned, it is performed using Cartesian values and a Proportional and Derivative (PD) approach. Next, the forces applied to the master are obtained following (1), where K_p and K_d are the proportional and derivative gains of the loop respectively. $\vec{e}(t)$ is the error between the UR5e robot and Phantom Omni Bundle TCP and $\vec{f}(t)$ is the final force applied to the operator.

$$\vec{f}(t) = K_p \cdot \vec{e}(t) + K_d \cdot \frac{\partial \vec{e}(t)}{\partial t} \tag{1}$$

A triple loop approach is presented in Fig. 1, with a control loop on joint and Cartesian level to each robot, as well as the force feedback to the operator. Then, a last signal gives the operator visual information about the state of the robots. As for the variables in Fig. 1 and with i being 1 or 2, $\vec{x_{di}}$ is the desired position for the one of the robots, $\vec{q_{di}}$ represents the desired joint position one of the UR5e and $\vec{u_{di}}$ is the torque command that contains all joint references for one of them. In addition, $\vec{q_i}$ and $\vec{x_i}$ are the current joint and Cartesian positions of one manipulator. Lastly, $\vec{f_i}$ is the force applied to one of the masters and $\vec{\tilde{f_i}}$ is the feedback force applied to the operator and calculated with (1).

All joint and TCP position information is recorded after performing a trajectory of dual object manipulation with both robots, so it can be used later to apply RLfD in a RL environment. With this information, the demonstrations provided will be set as a reference point to explore the horizon of possible manipulations, generalising from them and avoiding to start the RL training without any other information. In addition, it is possible to use traditional samples from a RL algorithm so other possible trajectories may be evaluated, considering other configurations. Moreover, trajectories performed by an expert can also be used as an evaluation sequence during the training episodes and, sometimes, they may take control of the process to correct deviations from desired behaviours [18].

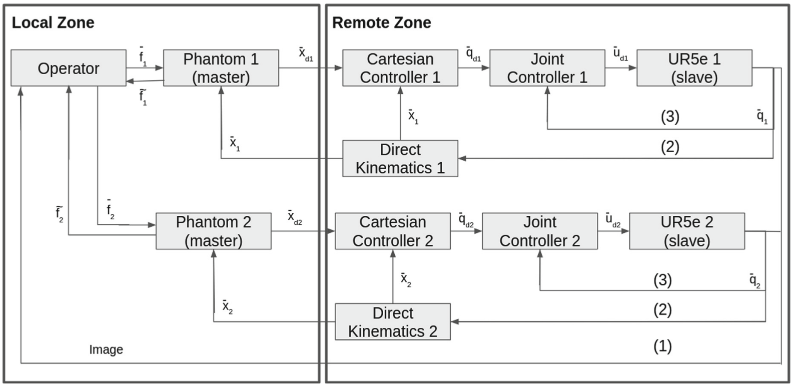


Fig. 1. Control loop. From outside to inside: (1) Visual feedback loop, (2) Cartesian position loop for one of the robots, (3) joint position loop for one of the robots.

The Gym platform offers an architecture based on actions and observations in order to perform all kind of RL, which is commonly used along with PyBullet direct simulations. For this reason, a learning environment has been designed and constructed to generate movement samples to be evaluated by the agent, who will follow a policy based on the information recorded by the dual-robotic haptic interface mentioned. The trajectories recorded and used to train the agent involve joint information in position, velocity and torque from the robots and grippers, as well as TCP Cartesian position and velocity in the world.

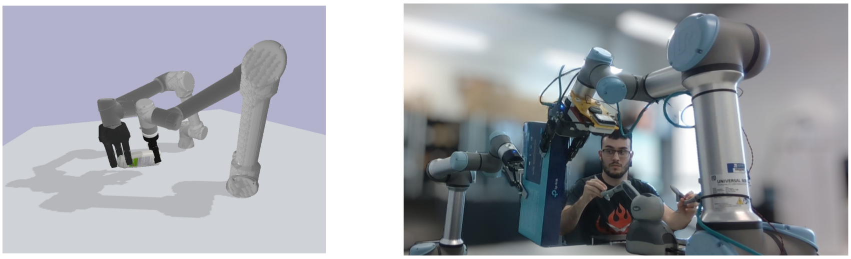
4 Experiments

In order to prove the quality of the obtained data involving positional trajectory and haptic information, some experiments and tests were conducted aiming towards this objective. In addition, comparison with a traditional teleoperation system is presented following the same goal.

4.1 Grasping Experiments

Both, real and simulation experiments were carried out in order to test the correct system performance in control, as well as in manipulation tasks, as it is represented in Fig. 2.

In Fig. 2a both robots are grasping a bleach cleanser from the Yale-CMU-Berkeley (YCB) [2] object dataset in a simulation environment generated by using PyBullet. The manipulation is effectuated at the same time in two areas of the object (top and bottom) laying down in a surface, allowing the operator to manipulate it coordinating both robots. On the other hand, in Fig. 2b the system is manipulating a box in a collaborative way, as both robotic arms are taking part in the process grasping the object at the same time. Moreover, the system permits passing objects from one robot to another if the operator is dexterous enough.



(a)

(b)

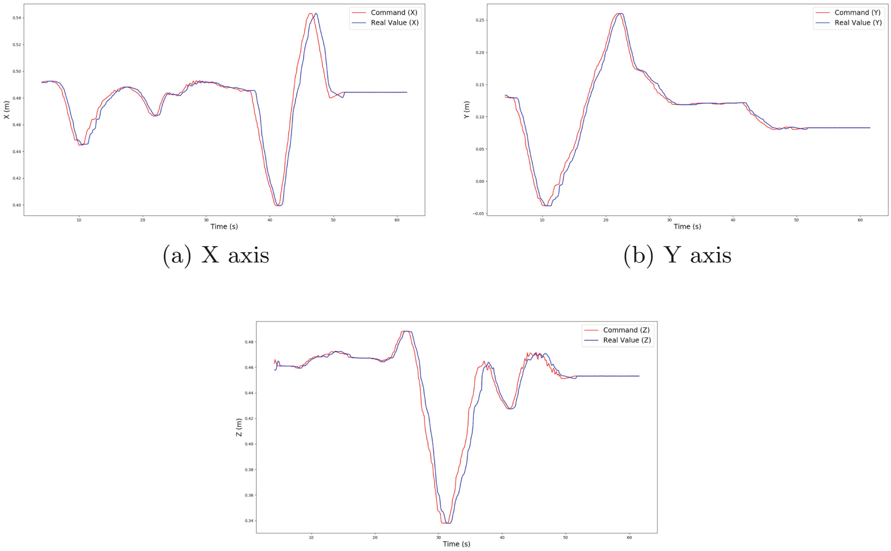
Fig. 2. Collaborative Manipulation experiments in simulation (a) and real environments (b) using the same architecture on each case.

4.2 Data Acquisition Experiments

Alternatively to the grasps experiments and to demonstrate the correct performance of the system in real environments, several trajectories were tested and recorded for later analysis. The goal is to prove that haptic information improves the generation of trajectory samples, resulting in more precise and smoother movements. Figure 3 show the movement of one of the robots on the three global axis taking as reference the UR5e's base frame while applying force feedback to the expert. Figure 3 show that robot's TCP follows the master commands with low error and oscillation, allowing the system to generate data of higher quality and more reliable than other approaches, such as kinematic manipulation demonstrations from a human or traditional data generation in RL training.

In addition, some other experiments were conducted using the same system but without applying haptic feedback. This way, the performance between data generation with and without effort information provided to the operator can be evaluated. In Fig. 3 it is shown a similar trajectory to Fig. 4 without applying haptic feedback.

Comparing the trajectories represented in Figs. 3 and 4 it can be concluded that the second one presents an increase in the error performing the trajectories

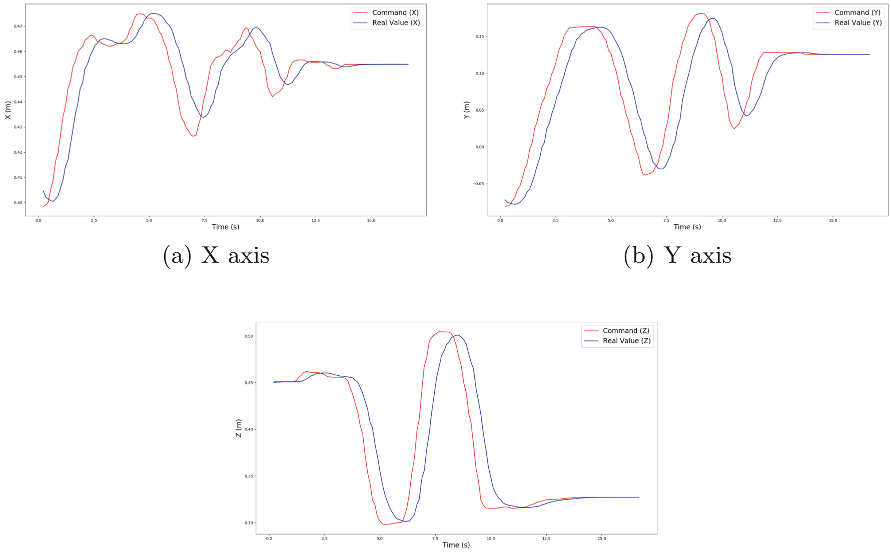


(a) X axis

(b) Y axis

(c) Z axis

Fig. 3. Reference (red) and output (blue) signals from the XYZ axis in a teleoperated trajectory using the haptic feedback in real environments.



(a) X axis

(b) Y axis

(c) Z axis

Fig. 4. Reference (red) and output (blue) signals from the XYZ axis in a teleoperated trajectory without haptic feedback in real environments.

Table 1. Evaluation of the system performance with and without haptic loop.

	Max. Error (m)	Min. Error (m)	Mean Error (m)	STD (m)
With Haptic	0.026	2e-05	0.00945	0.007
Without Haptic	0.16	8e-04	0.043	0.026

at first sight. Table 1 quantifies the performance of both cases; with and without applying haptic loop to the teleoperation.

From Table 1 it can be extracted that the data acquired from the system using haptic loop presents around 50% less error than the same system not applying effort feedback to the operator. As a consequence, trajectories generated with this dual-robot haptic interface reflect accurately the desired manipulation performed by the human, resulting in better results when leading the robot's TCP. Moreover, augmented data from trajectories such as in Fig. 3b for example, will have higher quality and be more reliable than the same demonstration taken from Fig. 4b.

5 Conclusion

In summary, a dual-robot haptic interface has been built to generate samples for future RLfD in the field of multiple robot control and object manipulation involving more than one agent. Moreover, the experiments conducted in real and simulation environments prove that force feedback improves the performance of teleoperation by giving the operator precise internal and external information about the robot's state and its surroundings.

This approach will cause the system to converge faster to a stable behaviour. Moreover, manipulation demonstrations from classic RL can be generated to explore the horizon of possibilities and generalize all possible trajectories that the system can produce once the training is finished.

6 Future Works

The purpose of this research is to have a dual-robot haptic interface designed for two robots to perform RLfD with dual robotic arms so as they can work together in multiple applications. The obtained trajectories will provide positive demonstrations when establishing innovative learning policies to be achieved by learning agents. The trajectories from the dual-robot haptic interface will be used to generate more trajectories, so the training with RL approaches becomes more effective than using random examples. In addition, expert demonstrations will be used as a reference to evaluate the robot's performance while training, allowing online evaluation without needing of a specific evaluation function.

This training will be aimed towards collaborative tasks like collaborative movement and dual manipulation, not only between robots but including humans sharing the same environment.

Acknowledgement. Research work was funded by the Spanish Government through the project PID2021-122685OB-I00.

References

1. Argall, B. D., Chernova, S., Veloso, M., Browning, B.: A survey of robot learning from demonstration. In: Robotics and Autonomous Systems, vol. 57, pp. 469-483. Elsevier (2009). https://doi.org/10.1016/j.robot.2008.10.024
2. Calli, B., et al.: Yale-CMU-Berkeley dataset for robotic manipulation research. Int. J. Robot. Res. **36**, 261–268. SAGE Publications Sage UK (2017). https://doi.org/10.1177/0278364917700714
3. Chen, Q., Dallas, E., Shahverdi, P., Korneder, J., Rawashdeh, O. A., Geoffrey Louie, W. -Y.: A sample efficiency improved method via hierarchical reinforcement learning networks. In: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 1498-1505. IEEE (2022). https://doi.org/10.1109/RO-MAN53752.2022.9900738
4. Clark, J. P., Lentini, G., Barontini, F., Catalano, M. G., Bianchi, M., O'Malley, M. K.: On the role of wearable haptics for force feedback in teleimpedance control for dual-arm robotic teleoperation. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 5187-5193. IEEE (2019). https://doi.org/10.1109/ICRA.2019.8793652
5. Corke, P. I.: A simple and systematic approach to assigning Denavit–Hartenberg parameters. IEEE Trans. Robot. **23**, 590–594. IEEE (2007). https://doi.org/10.1109/TRO.2007.896765
6. De Coninck, E., Verbelen, T., Van Molle, P., Simoens, P., Dhoedt, B.: Learning robots to grasp by demonstration. Robot. Auton. Syst. **127**, 103474. Elsevier (2020). https://doi.org/10.1016/j.robot.2020.103474
7. Gao, Q., Ju, Z., Chen, Y., Wang, Q., Zhao, Y., Lai, S.: Parallel dual-hand detection by using hand and body features for robot teleoperation. IEEE Trans. Hum.-Mach. Syst. **53**(2), 417-426. IEEE (2023). https://doi.org/10.1109/THMS.2023.3243774
8. Girbés-Juan, V., Schettino, V., Demiris, Y., Tornero, J.: Haptic and visual feedback assistance for dual-arm robot teleoperation in surface conditioning tasks. IEEE Trans. Haptics **14**(1), 44–56. IEEE (2021). https://doi.org/10.1109/TOH.2020.3004388
9. Girbés-Juan, V., Schettino, V., Gracia, L., Solanes, J.E., Demiris, Y., Tornero, J.: Combining haptics and inertial motion capture to enhance remote control of a dual-arm robot. J. Multimodal User Interfaces **16**, 219–238 (2022). https://doi.org/10.1007/s12193-021-00386-8
10. Hu, H., Zhao, Z., Yang, X., Lou, Y.: A Learning from Demonstration Method for Robotic Assembly with a Dual-Sub-6-DoF Parallel Robot. In: 2021 WRC Symposium on Advanced Robotics and Automation (WRC SARA), pp. 73-78. IEEE (2021). https://doi.org/10.1109/WRCSARA53879.2021.9612676
11. Arulkumaran, K., Deisenroth, M.P., Brundage, M., Bharath, A.A.: Deep reinforcement learning: a brief survey. IEEE Signal Process. Mag. **34**(6), 26–38. IEEE (2017). https://doi.org/10.1109/MSP.2017.2743240
12. Kyrarini, M., Zheng, Q., Haseeb, M. A., Gräser, A.: Robot learning of assistive manipulation tasks by demonstration via head gesture-based interface. In: 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), pp. 1139–1146. IEEE (2019). https://doi.org/10.1109/ICORR.2019.8779379

13. Laghi, M., et al.: Shared-autonomy control for intuitive bimanual telemanipulation. In: 2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids), pp. 1-9. IEEE (2019). https://doi.org/10.1109/HUMANOIDS.2018.8625047
14. Li, Y.: Deep reinforcement learning: an overview. In: arXiv preprint arXiv:1701.07274 (2017)
15. Lindner, T., Milecki, A.: Reinforcement learning-based algorithm to avoid obstacles by the anthropomorphic robotic arm. Appl. Sci. (2022). https://doi.org/10.3390/app12136629
16. Love, L.J., Book, W.J.: Force reflecting teleoperation with adaptive impedance control. In: IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 34, pp. 159–165. IEEE (2004). https://doi.org/10.1109/TSMCB.2003.811756
17. Nair, A., McGrew, B., Andrychowicz, M., Zaremba, W., Abbeel, P.: Overcoming exploration in reinforcement learning with demonstrations. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 6292–6299 (2018). https://doi.org/10.1109/ICRA.2018.8463162
18. Ramírez, J., Yu, W., Perrusquía, A.: Model-free reinforcement learning from expert demonstrations: a survey. Artif. Intell. Rev. **55**, 3213–3241. Springer (2022). https://doi.org/10.1007/s10462-021-10085-1
19. Ravichandar, H., Polydoros, A. S., Chernova, S., Billard, A.: Recent advances in robot learning from demonstration. In: Annual Review of Control, Robotics, and Autonomous Systems, pp. 297–330. Annual Reviews (2020). https://doi.org/10.1146/annurev-control-100819-063206
20. Sasaki, M., Muguro, J., Kitano, F., Njeri, W., Matsushita, K.: Sim-real mapping of an image-based robot arm controller using deep reinforcement learning. Appl. Sci. (2022). https://doi.org/10.3390/app122010277
21. Selvaggio, M., Abi-Farraj, F., Pacchierotti, C., Giordano, P. R., Siciliano, B.: Haptic-based shared-control methods for a dual-arm system. IEEE Robot. Autom. Lett. **3**(4), 4249–4256. IEEE (2018). https://doi.org/10.1109/LRA.2018.2864353
22. Si, W., Wang, N., Yang, C.: A review on manipulation skill acquisition through teleoperation-based learning from demonstration. Cogn. Comput. Syst. **3**, 1–16. Wiley Online Library (2021). https://doi.org/10.1049/ccs2.12005
23. Sun, X., Li, J., Kovalenko, A. V., Feng, W., Ou, Y.: Integrating reinforcement learning and learning from demonstrations to learn nonprehensile manipulation. IEEE Trans. Autom. Sci. Eng. **20**(3), 1735–1744. IEEE (2023). https://doi.org/10.1109/TASE.2022.3185071
24. Turlapati, S. H., Campolo, D.: Towards haptic-based dual-arm manipulation. Sensors **23**, 376. MDPI (2022). https://doi.org/10.3390/s23010376
25. Wu, R., Zhang, H., Peng, T., Fu, L., Zhao, J.: Variable impedance interaction and demonstration interface design based on measurement of arm muscle co-activation for demonstration learning. In: Biomedical Signal Processing and Control, pp. 8–18. Elsevier (2019). https://doi.org/10.1016/j.bspc.2019.02.008
26. Xie, Z.W., Zhang, Q., Jiang, Z.N., Liu, H.: Robot learning from demonstration for path planning: a review. Sci. China Technol. Sci. **63**(8), 1325–1334 (2020). https://doi.org/10.1007/s11431-020-1648-4
27. Zhang, Y., Zhao, X., Tao, B., Ding, H.: Multi-objective synchronization control for dual-robot interactive cooperation using nonlinear model predictive policy. IEEE Trans. Ind. Electron. **70**, 582–593. IEEE (2022). https://doi.org/10.1109/TIE.2022.3150090

28. Zhang, S., Xia, Q., Chen, M., Cheng, S.: Multi-objective optimal trajectory planning for robotic arms using deep reinforcement learning. Sensors (2023). https://doi.org/10.3390/s23135974
29. Zhao, W., Queralta, J. P., Westerlund, T.: Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 737–744, IEEE (2020). https://doi.org/10.1109/SSCI47803.2020.9308468