

Augmented Reality Platform for Collaborative Human-Robot Object Transportation Tasks

Graziano Carriero¹, Monica Sileo², Francesco Pierri², Fabrizio Caccavale¹

Abstract—This paper presents an Augmented Reality platform enabling bi-directional interaction between a human operator and a robot during collaborative object transportation tasks. The proposed system supports two interaction modes. In the first one, based on the leader-follower paradigm, the operator drives the motion while the robot assists the task by complying with the forces and torques applied during transportation. In the second mode, the operator specifies a target pose for the object via the Augmented Reality platform and the robot autonomously calculates its end-effector reference trajectory while maintaining compliant behavior to safely react to the wrenches exerted by the operator and/or the environment. Experimental results employing the Franka Research 3 collaborative robot and the Microsoft HoloLens 2 device demonstrate the robustness of the framework and its capability to assist the operator in transporting large objects while improving the intuitiveness of human-robot interaction.

I. INTRODUCTION

Human-Robot Collaboration (HRC) is widely implemented by scheduling sequences of tasks wherein operators and robots work toward a shared objective [1], [2]. This synergy leverages the complementary strengths of both agents: the robot's physical power generation capability and the human's cognitive abilities. For instance, humans' proactive capabilities can be exploited to oversee and optimize the physical power of robots when facing diverse task demands and unexpected disturbances. Consequently, achieving seamless integration between humans and robots represents a challenge of paramount importance in HRC.

In collaborative scenarios, the information stream can be *uni-directional*, proceeding from the robot to the human operator or vice-versa, or *bi-directional*, where a reciprocal communication channel is established. As detailed in [3], Human-Robot Collaboration interfaces aim at establishing a bi-directional information flow, which can be primarily classified into feed-forward and feedback modalities. Feed-forward interfaces manage the communication stream directed *to* the robot, originating from the human operator or the external environment [4]. A prominent example is provided by visual interfaces, which assist the user in the intuitive generation and

validation of trajectories [5]. Moreover, in contexts involving physical Human-Robot Collaboration (pHRC), haptic devices assume a crucial role by integrating the human in the loop. They facilitate a bi-directional signal stream, dealing with the force exchange, which is fundamental for tasks such as the co-transportation of objects [6]. Conversely, feedback interfaces operate in the opposite direction, i.e., *from* robot to human, aiming to augment the operator's perception of the task and its environment. Mixed Reality (AR) devices, for instance, are particularly well-suited for this purpose, enhancing human awareness through the overlaying of contextual information supplied by the robot [7], [8], [9].

In the field of pHRI, safety, dependability and performance must be integrated within a unified framework of optimality. Such an approach aims to guide future advancements in the design of robotic systems that can operate safely and reliably in human-centered environments, according to the vision of Industry 5.0 [10], [11], [12]. This research direction is motivated by the growing deployment of robots for object transportation tasks in industrial contexts, as well as in other domains such as service robotics and teleoperation. Such collaborative schemes enable the distribution of interaction forces among multiple robots, reduce the physical workload on humans and facilitate the manipulation of large or heavy objects while maintaining safe operating conditions.

The dynamic allocation of roles between human and robot agents is explored in [13] through the introduction of three distinct schemes for task effort distribution. A central concern in this framework is the robotic system's capacity to autonomously and appropriately adapt its own role during collaboration, which represents a necessity dictated by task dynamics, environmental uncertainty and unpredicted variations in human intent.

A variable impedance control is implemented in [14] for HR collaborative transportation tasks, where three control modes are defined: robot-dominant, human-dominant and collaboration mode. A learning mechanism for parameter estimation is employed for enhancing safety, by actively adapting stiffness and damping throughout the human-robot interaction. Similarly, the work in [15] proposes a variable admittance control approach for the HR object manipulation. In this scheme, the admittance parameters are dynamically adapted based on the desired trajectory and the external forces sensed by the robot, enabling the system to effectively respond to potential changes of the manipulated object. This approach further incorporates robust inverse dynamics control to stabilize the motion control laws, ensuring accurate trajectory tracking even in the presence of uncertainties.

¹ Department of Health Sciences, University of Basilicata, 85100 Potenza, Italy. {graziano.carriero, fabrizio.caccavale}@unibas.it

² Department of Engineering, University of Basilicata, 85100 Potenza, Italy. {monica.sileo, francesco.pierri}@unibas.it
Corresponding author: Graziano Carriero.

This research was funded by the Next Generation EU-Italian National Recovery and Resilience Plan (NRRP), Mission 4, Component 1, under the program PRIN 2022 PNRR of the Italian Ministry of University and Research. Project MELODY (Multi robot collaborative manipulation supporting Disassembly tasks) CUP C53D23008320001.

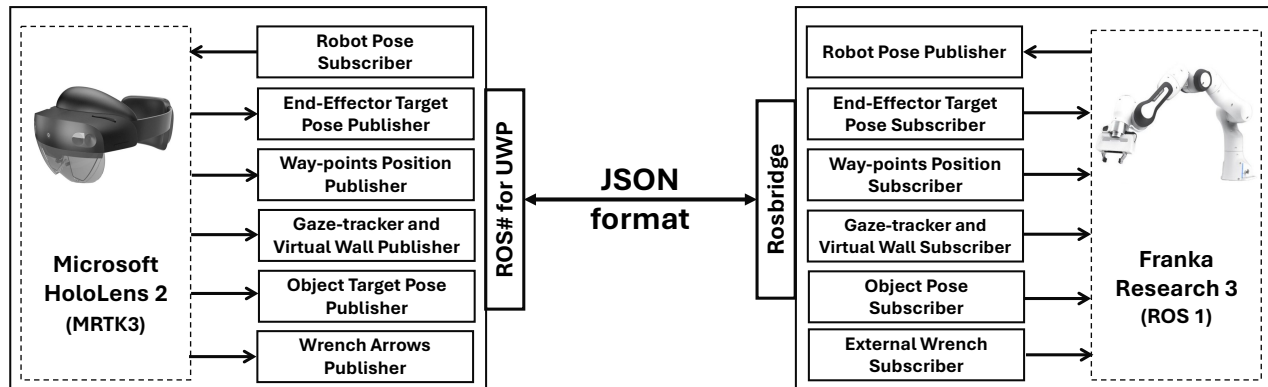


Fig. 1. System architecture: the Microsoft HoloLens 2 device, on the left; the Franka Research 3 robot, on the right. The communication stream is guaranteed via ROSBridge, employing JSON format messages.

Another critical issue arises from the translation–rotation ambiguity resulting from the coupling of multiple degrees of freedom in the object’s dynamics. The study in [16] addresses the difficulties encountered by human operators when applying large torques during collaborative object handling with a robot. By imposing a virtual nonholonomic constraint on the object, the robot’s motion is effectively restricted, thereby preventing object’s sideslip and enhancing the stability of the cooperative manipulation.

To address this challenge, further efforts have focused on the development of passive robotic systems, whose mechanical design specifically aims to mitigate the problem, allowing for efficient transmission of wrenches and smooth trajectory tracking. In [17], a system consisting of multiple mobile collaborative robots has been developed to enhance payload manipulability. In this case, the human operator exerts direct control over the trajectory and forces, acting as the primary guiding force to manipulate the object. Passive robotics is also explored in [18], where a passive robotic porter, comprising three omnidirectional wheels equipped with servo brakes and a control unit, is developed to assist humans in object co-manipulation. The additional passive joint is exploited to adjust the object’s orientation, while its position is regulated through the servo brakes.

II. MOTIVATION AND CONTRIBUTION

As previously discussed, Augmented Reality (AR) applications represent effective interfaces for Human-Robot Interaction in collaborative transportation contexts. They provide a bi-directional information flow that enables full integration of the human operator into the control loop.

The development of HR platforms for collaborative transportation is part of an advanced project aimed at creating a general-purpose augmented reality framework on the HoloLens 2, entitled *HumanTIX* [19]. The platform has been deeply enriched and developed so that the system is capable of addressing the integration of the human operator as an agent

in the loop, exploring dynamic role allocation and intentions exchange issues, previously investigated. It has been designed to ensure safe interaction between robots and low-skilled operators during tasks, such as robot trajectory planning, pick-and-places (see Section III) and object transportation (see Section IV). As assessed in [19], this approach reduces the physical workload associated with robot interaction, at the cost of a slight increase in cognitive load for novice users. The main aspect requiring attention concerns the potential information overload during the user’s first interactions. However, an appropriately configured level of robot autonomy can ensure a simple and effective user experience. Moreover, the platform implements novel features that address the need to incorporate safety guidelines by introducing the tracking of parameters and signals from both the human and the robot, with the aim of strengthening the bi-directional communication [12].

The application proposed in this work incorporates both feed-forward and feedback interaction modes (see Section IV). On one hand, it allows for the generation of the desired object trajectory; on the other hand, it enhances the operator’s perception by providing information about the forces exerted by the robot on the object. In this heterogeneous system, composed of a human operator, a collaborative robot and an AR device, the operator can dynamically adjust the robot’s level of autonomy, switching from a simple leader–follower scheme to a more advanced configuration in which the object trajectory is shared with the manipulator. In the latter case, the implementation of an admittance controller ensures the reduction of interaction forces and guarantees a safe collaboration (see Section V). Additional AR functionalities, e.g., controlling the opening and closing of the robot’s end-effector, are provided through an intuitive virtual control panel. Finally, thanks to its modular design, the application naturally supports the development and integration of new features that exploit the full capabilities of the augmented reality device (e.g., human gaze or hand movements tracking).

III. AUGMENTED REALITY FEATURES

The system architecture is illustrated in Figure 1. On the left, the Microsoft HoloLens 2 wearable device runs the Unity platform (v2021.3.22f1, 64-bit) [20], integrated with the Mixed Reality Toolkit v3 (MRTK3) [21]. This library provides essential components for developing cross-platform features within the Unity environment. On the right, the Robot Operating System (ROS) is in charge of managing the inputs and commands for the Franka Research 3 collaborative robot. Bi-directional communication between these two components is achieved through the exchange of JSON-formatted messages using the ROSBridge protocol, which enables non-ROS clients to interact with ROS systems via WebSockets.

This platform is built upon previous advancements developed in [12], [19], which provides:

- the robot URDF-hologram placement in the AR environment and calibration error assessment (about 5 mm in position, < 1 deg in orientation);
- the visual feedback of the manipulator's joint states through color-mapped virtual spheres and reachable workspace;
- the trajectory path generation via the interaction with virtual entities, including the end-effector hologram and the path way-points;
- the robot planned trajectory simulation and execution;
- the human gaze-tracker, for monitoring the operator concentration, and the virtual wall construction, for separating the human and robot workspaces.

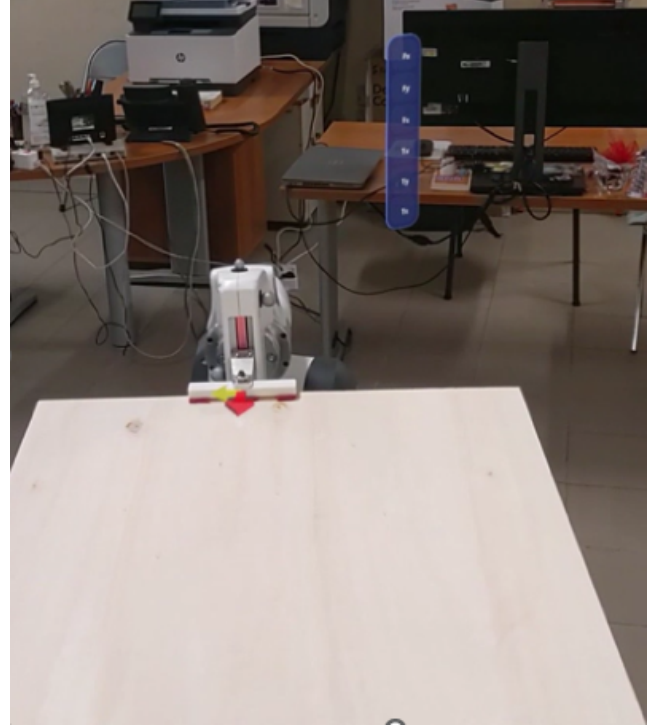
Additional AR features have been implemented for executing a collaborative object transportation task, including the introduction of virtual buttons for managing the end-effector status, calibrating the manipulated object and tuning the system's autonomy level through various collaboration modalities, which will be comprehensively discussed in Section IV.

IV. HUMAN-ROBOT COLLABORATIVE OBJECT TRANSPORTATION TASK

By leveraging the bi-directional communication established through the AR platform *HumanTIX*, the human operator can be seamlessly integrated as an *agent in the loop* within the control architecture. On the one hand, the human operator can exploit the information provided by the Augmented Reality platform, visualized through the HoloLens 2 display, to monitor the robot's current state in real time, which shows also a feedback of the interaction forces exerted on the object during the task execution. On the other hand, the operator can plan the object trajectory and suitably modify it to minimize the forces applied on the object or for collision avoidance. Furthermore, the operator is able to proactively update the task sequence, such as grasping, pick-and-place, or motion actions, according to the task requirements. Finally, the proposed framework allows the user to switch between different control modes, ranging from a leader-follower scheme to an object admittance control. In this way, the operator can exploit her/his superior cognitive capabilities to assist the robot in complex decision-making or adaptation processes, while the manipulator assist



(a) Third-person perspective.



(b) Human operator wearing the HoloLens device perspective.

Fig. 2. Two snapshots of the video showing the collaborative object transportation, available at <https://web.unibas.it/automatica/multimedia.html>.

her/him in physical demanding task, whose execution is beyond a single agent capability. A new 3D printed fingertip has been attached to the Franka Hand to ensure torque transmission to the robot and avoiding object side-slipping throughout the transportation task. Figure 2 shows two snapshots of the

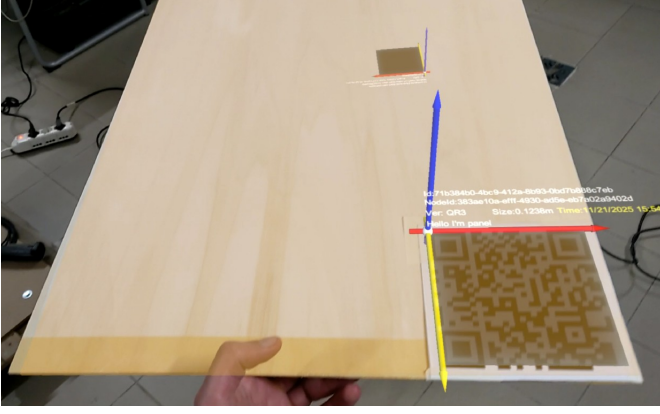


Fig. 3. The QR code used for the object calibration. Once detected, the virtual panel is superimposed on the real one.

collaborative object transportation task from an external perspective (Fig. 2a) and from the operator wearing the HoloLens device (Fig. 2b). The operator is provided with the wrenches exerted on the object, which are represented by virtual arrows whose size varies according to the numerical value of the wrenches, and can activate/deactivate the visualization of each component according to her/his needs. More in detail, the two interaction modalities for HR collaboration will be explored in the following subsections.

A. Leader-follower modality

In the first control mode, the human acts as the leader while the robot operates as the follower. In this configuration, the robot assists the operator in physically demanding tasks, such as object transportation, by exhibiting a compliant behavior. Specifically, the operator commands the desired object trajectory by manually guiding it within the shared workspace: she/he can tackle unexpected situations and provide fine adjustments of the trajectory of the grasped object at each time instant. On the contrary, the robot provides support in lifting and moving the object through compliant control. In this scenario, the AR interface helps the human operator grasp the object and provides visual feedback on the forces exerted by the robot on it in order to reduce its internal stress.

According to this modality, the robot behaves as a passive agent, merely reacting to the operator's actions. It is compliant with the wrenches sensed throughout the object motion, according to

$$M\ddot{x} + D\dot{x} = \gamma, \quad (1)$$

where $x = [p^T, \phi^T]^T$ is the robot end-effector pose, with ϕ denoting the Euler angles triplet and M and D are the gain matrices for the robot virtual mass and damping, respectively. Moreover, $\gamma = T_A^T(\phi)h_e$, with

$$T_A(\phi) = \begin{bmatrix} I_3 & O_3 \\ O_3 & T(\phi) \end{bmatrix}, \quad (2)$$

where $T(\phi)$ relates the Euler angles derivative $\dot{\phi}$ to the angular velocities ω and h_e is the wrench measured at the

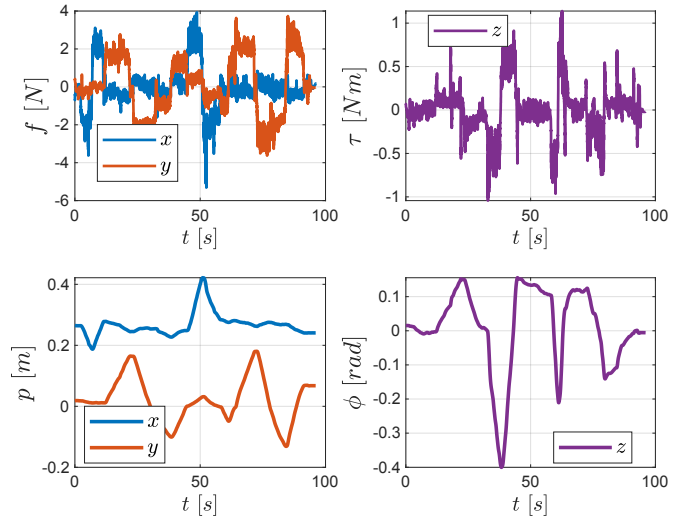


Fig. 4. Robot trajectory (bottom) and corresponding filtered sensed wrench (top).

manipulators' end-effectors. Notice that no spring term is included in (1), since the desired end-effector pose, x_d , can be considered as the actual pose of the manipulator in every time instant.

B. Admittance control modality

In the second control mode, the human operator still defines the object trajectory, but this information is fed into the robot controller via the Augmented Reality interface. As in the previous case, the AR platform assists the operator during the initial grasping phase and provides real-time visualization of the robot state and the applied wrenches.

After performing the calibration of the manipulator, the AR device identifies the initial pose of the object through the detection of the QR code placed on it (see Fig. 3). Consequently, the virtual object reference frame is displayed within the AR scene, thereby the user can translate and rotate it to assign the target pose of the object. Based on this information, a path is generated in the task space and the robot exploits the closed-chain constraints [22] to compute the corresponding desired end-effector trajectory ($x_d, \dot{x}_d, \ddot{x}_d$).

Throughout the task execution, an admittance control scheme is used to reduce the wrenches exerted on the object. The robot computes the human wrench exerted on the object, along with the forces arising from an eventual contact with the external environment, $\gamma_o = T_A^T(\phi)h_o$, where $h_o = Gh_e$ and G represents the grasp matrix. Both these contributions determinate the wrench exerted at the object center of mass, according to

$$M_o\ddot{e}_o + D_o\dot{e}_o + K_o e_o = \gamma_o, \quad (3)$$

where $e_o = x_{o,d} - x_{o,r}$ is the object tracking error, with $x_{o,d}$ and $x_{o,r}$ denoting the object desired and reference pose, respectively. In this way, the manipulator acts as an active agent since it is in charge of limiting both external and

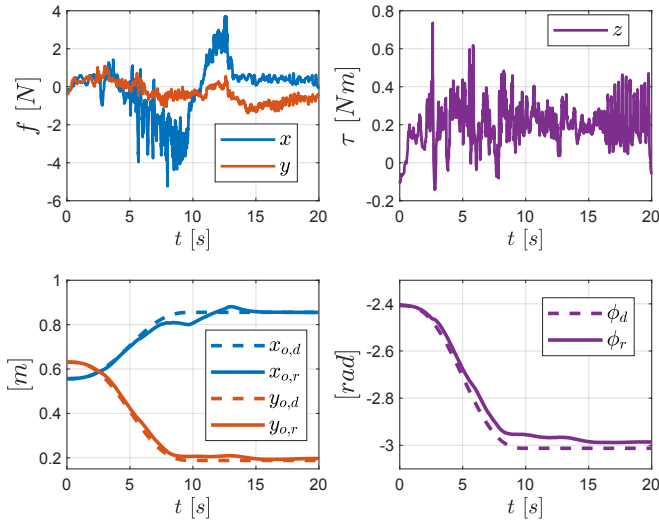


Fig. 5. Robot trajectory (bottom) and corresponding filtered sensed wrench (top). In the bottom row, dashed and lines represent the desired and the reference trajectories, respectively.

internal forces acting on the object, by means of the admittance control framework. On the other side, the human operator can change the robot trajectory by applying forces and torques to the object. This results in a modification of the end-effector reference trajectory, x_r , which the robot subsequently follows. When the operator stops to apply wrenches, the elastic term in (3) drives the system back to the desired trajectory, allowing the object to reach the predefined target pose.

Finally, the human operator can switch between the two proposed collaborative control methods through the AR interface, enabling intuitive and flexible task supervision and execution.

V. EXPERIMENTAL RESULTS

Two experiments have been carried out, one for each interaction mode. The workpiece to be transported is a wooden board of size $0.9 \times 0.45 \times 0.003$ m and weight 0.5 kg.

Figure 4 shows the interaction wrench applied to the object by the human operator and the corresponding variations in the object pose during the execution of the collaborative transportation task through the first interaction mode. In this experiment, that lasts approximately 100 s, the object motion is constrained in the xy plane. Thus, as can be seen in the two graphs in the top row of Fig. 4, only the forces along the x and y axes and the torque around the z axis are depicted. In the bottom row of the same figure, the x and y coordinates of the position and the z coordinate of the orientation for the end-effector are reported. As expected, the robot complies the sensed forces and torques, following the object trajectory imposed by the human operator.

In the second experiment, executed enabling the second interaction mode, the human operator moves the object digital twin in order to specify the target pose and commands the robot to start the motion by using the AR interface. Then, the robot computes the desired trajectory for its end-effector

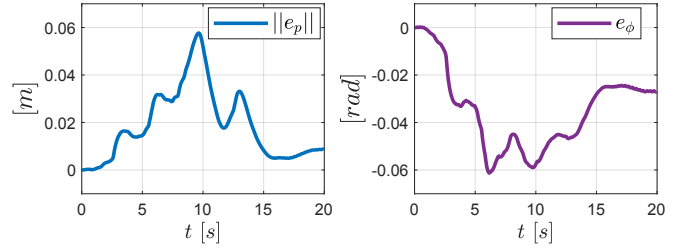


Fig. 6. Norm of the object position (left) and orientation (right) tracking errors.

using a third-order polynomial time law, which lasts 10 s. During the execution of the task, the robot tracks the desired trajectory while maintaining compliant behavior with respect to possible interaction forces and torques. In the first phase of the motion, the human operator also follows the trajectory as passive agent, then, around $t = 6$ s, starts to exert wrenches on the object. The robot immediately reacts, deviating from the desired trajectory according to the sensed wrench. This is shown in the bottom row of Fig. 5, where it can be seen that the reference trajectories (solid lines) deviate from the desired ones (dashed lines) according to forces and torques exerted by the operator. Around the time instant $t = 13$ s, the human operator stops to exert wrenches and, then, the robot resumes the desired trajectory and attempts to reach the target point by exploiting additional 10 s of steady-state time. Residual forces (along the y -axis) and torques (along the z -axis) measured by the robot can be observed in the top row of Fig. 5. As a consequence, the robot achieves the desired final pose with a residual tracking norm error of ~ 0.009 m and 0.025 rad, as shown in Fig. 6.

VI. CONCLUSIONS

This paper presents an Augmented Reality platform aimed at allowing bi-directional interaction between a human operator and a robot during object transportation tasks. The platform allows the human operator to use two interaction modes. The first one is based on the leader-follower paradigm, where the operator is in charge of imposing the motion and the robot has to assist her/him as passive agent. The second one forecasts that the robot has to follow a desired trajectory for the object (previously defined by the operator) and, at the same time, has to exhibit a compliant behavior in order to safety react to wrenches imposed by the human or the environment.

Experiments conducted with Microsoft HoloLens 2 and the collaborative Franka Research 3 robot confirm the robustness of the proposed approach and demonstrate the system's capability to assist the operator in transporting heavy objects, while also enabling more intuitive interaction with the robot through the virtual reality interface.

Future work will focus on integrating a mobile robotic platform to extend the operational workspace beyond the limits of a single manipulator during the transportation task.

REFERENCES

- [1] T. Li, S. Yu, B. Sun, Y. Li, X. Wang, Y. Pan, C. Song, Y. Ren, Z. Zhang, K. T. Grattan *et al.*, “Bioinspired claw-engaged and biolubricated swimming microrobots creating active retention in blood vessels,” *Science Advances*, vol. 9, no. 18, p. eadg4501, 2023.
- [2] H. Wang, S. Yu, J. Liao, X. Qing, D. Sun, F. Ji, W. Song, L. Wang, and T. Li, “A robot platform for highly efficient pollutant purification,” *Frontiers in Bioengineering and Biotechnology*, vol. 10, p. 903219, 2022.
- [3] A. Ajoudani, A. M. Zanchettin, S. Ivaldi, A. Albu-Schäffer, K. Kosuge, and O. Khatib, “Progress and prospects of the human-robot collaboration,” *Autonomous Robots*, vol. 42, 06 2018.
- [4] D. Sirintuna, A. Giammarino, and A. Ajoudani, “An object deformation-agnostic framework for human–robot collaborative transportation,” *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 2, pp. 1986–1999, 2024.
- [5] D. Sun, A. Kiselev, Q. Liao, T. Stoyanov, and A. Loutfi, “A new mixed-reality-based teleoperation system for telepresence and maneuverability enhancement,” *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 1, pp. 55–67, 2020.
- [6] L. Guo, Z. Lu, and L. Yao, “Human-machine interaction sensing technology based on hand gesture recognition: A review,” *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 4, pp. 300–309, 2021.
- [7] T. Chakraborti, S. Sreedharan, A. Kulkarni, and S. Kambhampati, “Projection-aware task planning and execution for human-in-the-loop operation of robots in a mixed-reality workspace,” in *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4476–4482.
- [8] S. Y. Gadre, E. Rosen, G. Chien, E. Phillips, S. Tellex, and G. Konidaris, “End-user robot programming using mixed reality,” in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 2707–2713.
- [9] M. Ostanin, S. Mikhel, A. Evlampiev, V. Skvortsova, and A. Klimchik, “Human-robot interaction for robotic manipulator programming in mixed reality,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 2805–2811.
- [10] A. De Santis, B. Siciliano, A. De Luca, and A. Bicchi, “An atlas of physical human–robot interaction,” *Mechanism and Machine Theory*, vol. 43, no. 3, pp. 253–270, 2008. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0094114X07000547>
- [11] A. Bicchi, M. Bavaro, G. Boccadamo, D. De Carli, R. Filippini, G. Grioli, M. Piccigallo, A. Rosi, R. Schiavi, S. Sen, and G. Tonietti, “Physical human-robot interaction: Dependability, safety, and performance,” in *2008 10th IEEE International Workshop on Advanced Motion Control*, 2008, pp. 9–14.
- [12] M. Sileo, G. Carriero, C. Brancato, N. Calzone, F. Pierri, F. Caccavale, and R. Mozzillo, “Safety features for humantix: An augmented reality platform to enhance human–robot collaboration,” in *Advances on Mechanics, Design Engineering and Manufacturing V*, C. Machado del Val, R. Miralbes Buil, G. Peris Fajarnés, M. Moncho Santonja, C. Rizzi, and L. Roucoules, Eds. Cham: Springer Nature Switzerland, 2025, pp. 557–572.
- [13] A. Mörtl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, “The role of roles: Physical cooperation between humans and robots,” *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1656–1674, 2012. [Online]. Available: <https://doi.org/10.1177/0278364912455366>
- [14] Y. Pan, Z. Li, T. Shi, and C. Wen, “Composite learning variable impedance robot control with stability and passivity guarantees,” *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 119–126, 2024.
- [15] M. Mujica, M. Crespo, M. Benoussaad, S. Junco, and J.-Y. Fourquet, “Robust variable admittance control for human–robot co-manipulation of objects with unknown load,” *Robotics and Computer-Integrated Manufacturing*, vol. 79, p. 102408, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0736584522000941>
- [16] H. Arai, T. Takubo, Y. Hayashibara, and K. Tanie, “Human-robot cooperative manipulation using a virtual nonholonomic constraint,” in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, vol. 4, 2000, pp. 4063–4069 vol.4.
- [17] M. L. Elwin, B. Strong, R. A. Freeman, and K. M. Lynch, “Human-multirobot collaborative mobile manipulation: The omnid mobocots,” *IEEE Robotics and Automation Letters*, vol. 8, no. 1, p. 376–383, Jan. 2023. [Online]. Available: <http://dx.doi.org/10.1109/LRA.2022.3226366>
- [18] Y. Hirata, H. Song, Z. Wang, and K. Kosuge, “Control of passive object handling robot with free joint for reducing human assistive force,” in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 1154–1159.
- [19] G. Carriero, N. Calzone, M. Sileo, F. Pierri, F. Caccavale, and R. Mozzillo, “Human-robot collaboration: An augmented reality toolkit for bi-directional interaction,” *Applied Sciences*, vol. 13, no. 20, 2023. [Online]. Available: <https://www.mdpi.com/2076-3417/13/20/11295>
- [20] Unity Technologies, “Unity,” 2023, game development platform. [Online]. Available: <https://unity.com/>
- [21] —, “Mixed reality toolkit 3,” 2023, mixed Reality Toolkit. [Online]. Available: <https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk3-overview/>
- [22] F. Caccavale, P. Chiacchio, A. Marino, and L. Villani, “Six-dof impedance control of dual-arm cooperative manipulators,” *IEEE/ASME Transactions On Mechatronics*, vol. 13, no. 5, pp. 576–586, 2008.