An Intuitive Augmented Reality Interface for Task Scheduling, Monitoring, and Work Performance Improvement in Human-Robot Collaboration

Alessandro De Franco[1,2], Edoardo Lamon[1,3], Pietro Balatti[1,3], Elena De Momi[2], and Arash Ajoudani[1]

***Abstract*—One of the open challenges in Human-Robot collaborative tasks is to provide a simple way for humans to understand robotic systems' plans and actions and to interact with them in the most natural way. Towards the direction of a natural interaction between human and robots, we propose a simple and intuitive interface exploiting the emergent Augmented Reality (AR) technology. This interface aims not only to enhance human awareness of the robot status and planned actions during collaborative tasks, but also to improve the quality of the work. The presented interface enables the human operators to interact with the system sending commands (gesture or vocal) and receive instant feedbacks (holograms and sound) through an AR device, enabling an intuitive way to coordinate human and robot actions. The interface validation is performed through two industrial scenario experiments, i.e., a collaborative assembly of a metallic structure and a collaborative polishing. The experimental results showed that the presented AR interface could be exploited to improve human operators' performance while performing an industrial task. At the same time the usage is also perceived as beneficial by ten subjects that tested the interface in the experiments.**

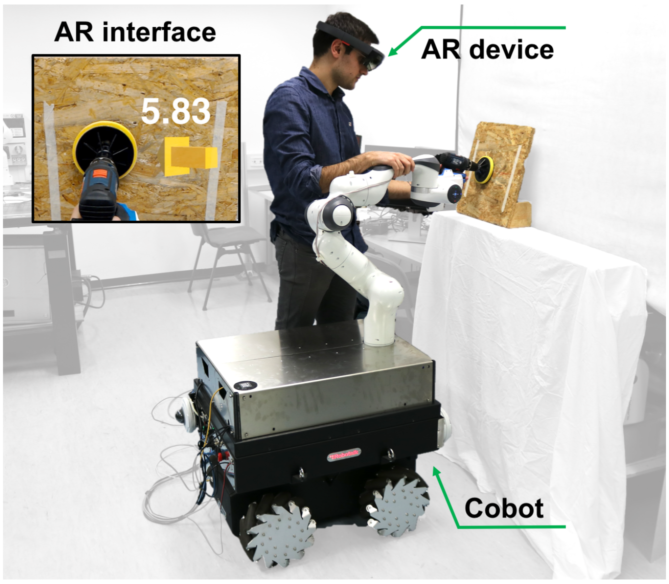

Fig. 1. Human-Robot collaborative task supported by an augmented reality interface. The human operator receives instant feedback about the applied force.

I. INTRODUCTION

The recent development of lightweight collaborative robots (cobots), in industrial manufacturing scenarios, depicts new prospects in the field of physical Human-Robot Interaction (pHRI). Nowadays, humans and cobots coexist within the same workspace without fences. Cobots, indeed, can be exploited in repetitive tasks where high-precision motion and considerable power capacity are required. The research interest in obtaining an efficient Human Robot Collaboration (HRC) is constantly growing [1]–[3]. Moreover, when dealing with humans, cobots should present also safety features. To ensure human safety in the workspace by avoiding injuries, robots must guarantee compliant behaviours, that might be the result of variable impedance control [4] and be able to detect collision and react accordingly [5]. Moreover, repetitive execution of heavy tasks could induce on the human side an excessive and continuous overloading of the body joints which may result in injuries and chronic diseases. Therefore, cobots should be exploited to reduce human body joint torques [6] or muscle fatigue [7]. In this way, the potential risk of injury related to the task is reduced.

The open challenge is to create a simple and intuitive system in which humans and cobots could interact in the most natural way. In human working teams, people communicate their intentions, besides spoken language, through gestures, facial expressions and, in general, with body motions. The absence of these rich and immediate communication means in human-robot teams decreases the trust of the human on the robot and the perception of robot user-friendliness, even if the environment is safe for the human. To face the challenge of a simple interface that allows human and robot to interact and coordinate, several devices have been developed, for instance, wearable and haptic devices [1].

A more recent trend in HRC consists in the introduction of augmented reality (AR) applications in industrial working environments. AR combines real-world and computer-generated data to enhance the user experience, differently from a purely virtual environment, in which the user is totally immersed. Several studies already show that AR improves HRC, in particular, improving task efficiency by helping workers to understand robot intent [8]. Furthermore, *Terenzi et al.* have demonstrated that AR application produces a reduction of

[1] HRI[2] Lab, Dept. of Advanced Robotics, Istituto Italiano di Tecnologia, Genova, Italy.
[2] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, Italy
[3] Robotics and Automation, Dept. of Information Engineering, Universita' degli Studi di Pisa, Pisa, Italy.
alessandro.defranco@iit.it

978-1-7281-0968-8/19/$31.00 ©2019 IEEE

costs up to 25% and an improvement in performance up to 30% in the manufacturing industry [9].

Other applications of AR in industrial scenarios have been presented. For example, Boeing workers exploit Google Glasses to help workers in the construction of aircraft wire harnesses [10]. To overcome the discomfort of the platform, other research introduces wearable devices that projects information on surfaces. *Funk et al.* proposed a system based on a projector that supports employees in assembly tasks by virtual information and instructions [11]. *Ruffaldi et al.* [12] developed a head-up display integrated into an industrial helmet that provides additional information about the workplace. An Android-based application for programming and monitoring industrial robots using a tablet is presented by *Mateo et al.* [13]. Even if the tablet presents a simple interface to communicate with the robot, workers cannot work and interact with the system at the same time. In the paper, indeed, a worker holds the table and another performs the collaborative task. In general, hand-free devices are more suitable for physical tasks in industrial environments.

In this paper, we propose a simple and intuitive augmented reality (AR) interface for human workers in collaborative tasks. The main purpose is to improve user experience, increasing worker confidence in the system and his awareness of the robot actions during collaborative tasks. We refer in particular to situations in which human and robot do not only coexist in the same space, but also interact between themselves and the environment. We will exploit the mixed reality smartglasses (Microsoft HoloLens), a device that allows not only the projection of holograms in a head-mounted display (HMD) but also presents new features like voice commands, as well as gestures recognition and gaze tracking, opening new prospects on the way human can interact and control robots.

In Sec. II we propose a flexible and customisable AR interface and its benefits in collaborative industrial tasks. To validate the effectiveness of the interface, we physically performed two different tasks using MOCA [14], a recently introduced versatile mobile-based assistive robot, that includes a Panda robotic arm by Franka Emika, equipped with the underactuated Pisa/IIT SoftHand [15], that are mounted on top of a SUMMIT-XL STEEL mobile platform by Robotnik. The tasks were composed of an assembly of aluminium profiles and a collaborative polishing. The first experiment shows how to exploit HoloLens features to make the human able to supervise and coordinate robot performances. The second task was performed by 10 subjects in different conditions and the performance was evaluated through physical data and through the results of a questionnaire. The results of the experiments are discussed in Sec. V. Finally, the conclusion of the work is presented in Sec. VI.

II. AUGMENTED REALITY INTERFACE FOR HUMAN-ROBOT COLLABORATIVE TASKS

To obtain the desired human-cobot collaboration, we would like to exploit human cognitive skills with cobot's physical power generation capacity. To do that, a real-time communication between the two agents is needed [16]. The AR interface, indeed, should be able to receive feedback from the smart work cell and eventually elaborate these data to obtain meaningful information for the human. In this way, workers will be more aware of the work cell's current status and future plans. On the other hand, human commands could improve both task scheduling, by creating an intuitive mean of task scheduling, and task performance by coordinating the two systems.

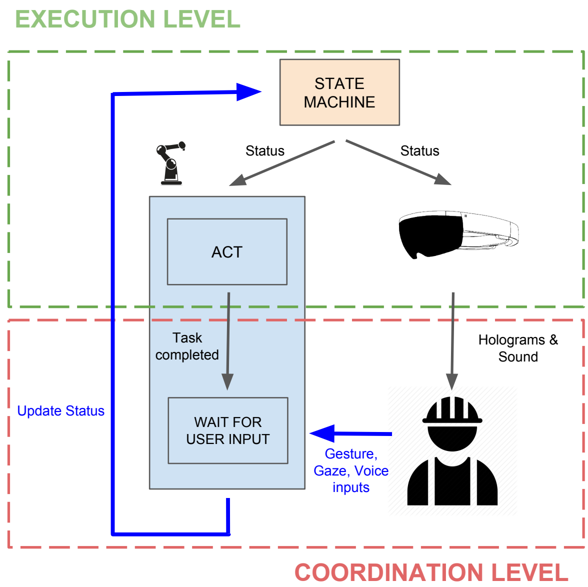

Fig. 2. State machine sends status information to the HoloLens and Robot, at the Execution level. Worker wearing HoloLens receives AR feedback (Holograms and Sounds) while Robot accomplishes the task. At the coordination level, robot switches in a break state, waiting for the workers' input. Human operator can trigger the system through a gesture, gaze or vocal command.

The AR device, i.e. Microsoft HoloLens, allows the human user to send inputs to the robotic system in different ways: vocal input, hand gesture input, and gaze. At the same time, the user could receive system information in the form of holograms (visual feedback) and sound (audio feedback). Although non-verbal interaction has already been introduced in HRC, as in the case of haptic cues [17], other devices are not able to provide a wide range of input/output data, streaming simultaneously audio and visual feedback while allowing the human user to interact in different ways.

The first goal of the presented AR interface is to help workers to interact in the most user-friendly way possible with the cobot. In this way, we tried to identify the type of data that can be the useful for the task execution, which should be streamed to the worker, either using the visual or audio feedback, and sent if requested. Another important aspect of the interface considers the overload of information. We want to provide the user with the smallest amount of information needed to achieve the task, not to stress or mentally overload them. Moreover, this additional data should not conflict with other environmental aspects of the real scene (e.g., occluding the workspace), to make the AR experience as less invasive as possible. The interface configuration depends on the particular task performed, but we will try to give a general overview. We will be more specific in the experimental section, Sec. IV. In general, the AR interface should provide instant feedback from the robotic system to the user, from high level information like the scheduled plan for an assembly task or current and future high-level cobot action, to lower-level system status data, in terms of Cartesian positions and velocity, force and torques applied to/from the end-effector, and even motor temperature of the cobot. These data are not just useful for safety reasons to improve workers awareness in the execution of the task, but also might be required for a correct task execution.

On the other hand, the user inputs can be used to coordinate cobot performances with the one of the workers. For instance, in order to give the cognitive lead of the task execution, we will allow the human worker to trigger the system status progress. After the performance of certain actions, the system will switch to a break state, waiting for the worker's input, that might be either a voice command or a gesture, according to the particular environment or task. The architecture of the proposed coordination system is depicted in Fig. 2.

A. Visual feedback

The AR interface communicates the cobot's status during the task, to make the user aware of the robot's movements. In addition, to raise awareness on the cobot's status, we added a light that communicates the current actions of the cobot:

- Green $\rightarrow$ robot moving away from the user.
- Red $\rightarrow$ robot approaching the user.
- Yellow $\rightarrow$ robot in waiting state.

When the system is waiting for human feedback (yellow light), the user can inform the system by sending commands via the AR device. The coloured lights are used as an effective non-verbal communication mode to help humans understand the robot's status and actions, similarly to the work of *Baraka et al.* [18].

B. Sound feedback

The users can understand what is happening in the environment by perceiving also sound feedback when the robot or the 3D-holograms are out of their field of view. The embedded speakers in the AR device allow the user to hear the simulation of the sounds coming from real physical objects (i.e. robot) or from 3D-holograms. These information feedback are useful, for example, in the interaction of the user with multiple robots that can communicate the beginning of each action through a proper sound.

C. User input

The choice of user inputs depends on the task to be performed. The gaze, for instance, estimated from the head orientation, can be used to place holograms in space or giving reference for the robot Cartesian trajectory. Another possible input is given by the vocal commands. So far, we have introduced the following vocal commands:

- STOP ROBOT: stops immediately the robot.
- NEXT ACTION: a hologram with the next robot action appears.
- COMMAND LIST: a hologram with the list of available vocal command appears.

Vocal commands guarantee users maximum handling ability but in the noisy industrial environment could not be effective. For this reason, our interface includes in the interface also the gesture command "Air Tap" (lift and flex the index finger for selecting an operation) exploiting the gesture recognition capability of the HoloLens platform.

III. EXPERIMENTAL SETUP

In this section, we will introduce the collaborative framework that we exploited in the experiments. The framework is composed by two main components (see Fig. 1): the AR device, i.e. the Microsoft HoloLens and a collaborative robot. In the experiments we used MOCA, a versatile mobile-based assistive robot. The HoloLens is a fully standalone HMD device that provides immersive AR. The device runs on Windows 10 and feature a Holographic Processing Unit and a 32-bit Intel CPU. The sensory system contained in the device hardware provides excellent stability of the 3D holograms in the real world. With Microsoft HoloLens, it is possible to place 3D-holograms into space through the environmental scanning. Moreover, the integrated gesture and speech recognition allow a real-time interaction with holograms. The gaze direction, used to place holograms in the environment, is estimated using the orientation and position of the user's head. Finally, the vocal input is possible because Hololens features four microphones, voice, and speech recognition.

The system architecture, in Fig. 3, allows double channel communication, from the robot to the HoloLens and viceversa, through User Datagram Protocol (UDP). UDP provides a real-time data transmission between the devices via WiFi. The robot control architecture is implemented using ROS. Microsoft provides different software to develop applications for Universal Windows Platforms. One of these is the Unity development software: a commercial cross-platform game creation system for creating interactive media. The Unity platform

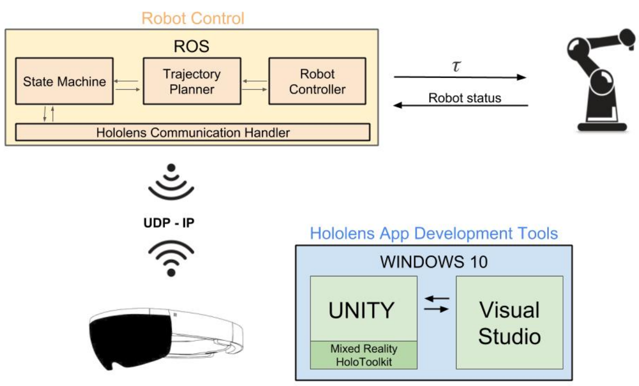

Fig. 3. System architecture enables double channel HoloLens-Robot communication via WiFi (UDP-IP). Hololens App Development Tools represent the environment in which a holographic 3D application is developed. The main components involved in the robot control are shown above.

provides many tools for holographic 3D development and simulation. In our case, it is used to create the AR scene in which the user is immersed. Moreover, we eploited Microsoft Mixed Reality Toolkit that is a Unity library containing scripts that allows the HoloLens user to interact with holograms. To develop and build C# scripts, Visual Studio was used. The result is a holographic 3D application that can be executed on the HoloLens.

IV. EXPERIMENTS

Two proof-of-concept experiments were performed in an industrial scenario, i.e., a collaborative assembly of aluminium profiles and a collaborative polishing of a rough surface. The main goal of these experiments is, on the one hand, to show how AR devices could be exploited to help the human operator in collaborative tasks, while, on the other hand, to evaluate if the human performance improves. To do that, the second experiment has been proposed with 3 different setups to 10 subjects and after the performance of the task with each setup a performance questionnaire has been presented.

A. Collaborative assembly

In the first experiment, we performed the collaborative assembly of a metallic structure composed by two aluminium profiles of different sizes, held together through a corner joint with screws and nuts. The high-level assembly task has been decomposed in a sequence of actions. A priori, these actions could be performed either from human worker or from the cobot (a Franka Emika Panda equipped with the standard gripper). The decomposition breaks down the task in a sequence of actions (single flow) and in parallel actions (multiple flows) and then assigns each action to an agent according to the algorithm presented in [19]. In single flow sequence of actions, the challenge consists in having a smart system able to recognize when each of the actions has been performed and the following action should start, while, in multiple flows of actions, to coordinate them when two different flows merge in the same one.

To simplify the operation of the worker, we exploit the HoloLens platform and the AR interface previously described. First, before the beginning of the task execution, the flow of action (task schedule) assigned to each agent (robot or human) has been projected on the HoloLens display. Second, during the execution of each action, the HoloLens projects also the current action performed both by human and robot. In this way, the human is always aware of the actual action of the cobot. Finally, we used the HoloLens gesture capture system to indicate to the system that the human has ended its current task and can move to the next one, as explained in Fig. 2.

First, the cobot picks-and-places small aluminium profile from the storage desk to the work bench. The activation of the robot was communicated to the worker also with a sound feedback (beep sound). During the experiment the system assembly status was shown to the worker through 3D holograms. Then the cobot picked the corner joint from the storage desk, approached it in assembly area and the system

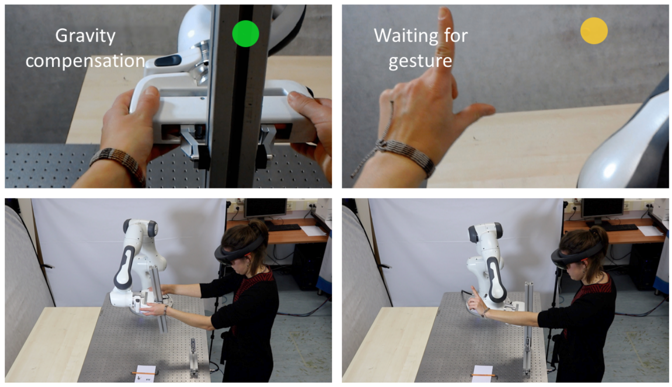

Fig. 4. Experimental collaborative assembly task (IV-A). On the left, the cobot is in gravity compensation mode and the operator can move it without any physical effort. On the right the system switches to a break state (waiting for an input) and the human operator is ready to send a gesture command. The user's view through HoloLens is shown in the pictures above: holograms communicate assembly status and the cobot's movement (see Sec. II-A).

switched to a break state, waiting for the worker input. After receiving the worker gesture input, the cobot hands over the corner joint. Afterwards the worker can attach the corner joint to the small profile: he inserts a nut and fastens it with a screw using an Allen key. After the cobot picked long aluminium profile and approached it in the assembly area, the cobot switched to gravity compensation control (Fig. 4). In this mode the cobot compensated for the weight of the long profile and the worker could align it to the corner joint without any physical effort. Worker attached the long aluminium profile to the small profile-joint corner structure, as done in the previous assembly task. Finally he sent an input through the gesture to the system, the cobot switched to the Cartesian impedance control and returned to the home position, where it is ready for a new assembly task.

B. Collaborative polishing

In this experiment the worker should perform the polishing of a rough wooden surface with a polisher, following a linear trajectory. The polisher is mounted on the end-effector of the robot through the Pisa/IIT SoftHand. The weight of the polisher is completely sustained by the robot thanks to the Cartesian impedance control with high impedance on the z-axis, but the operator can easily move the robot in the other 2 directions since the stiffness is almost null. The polishing task was performed by 10 subjects, between 25 and 35 years old, who were not aware of the purpose of the experiment. Subjects were required to keep the force applied perpendicular to the surface (on x-axis) equal to 10 Newtons approximately, for 30 seconds. The same experiment was performed in three different experimental settings in the following order: without any force feedback to the human operator, with a force feedback displayed on a screen, and the last with the force feedback projected on the HoloLens display. The latter interface illustrated to the user with two different holograms: one with a number representing the force expressed in Newton and another with an arrow that changes color (green if the applied force is close to the desired force, red if really

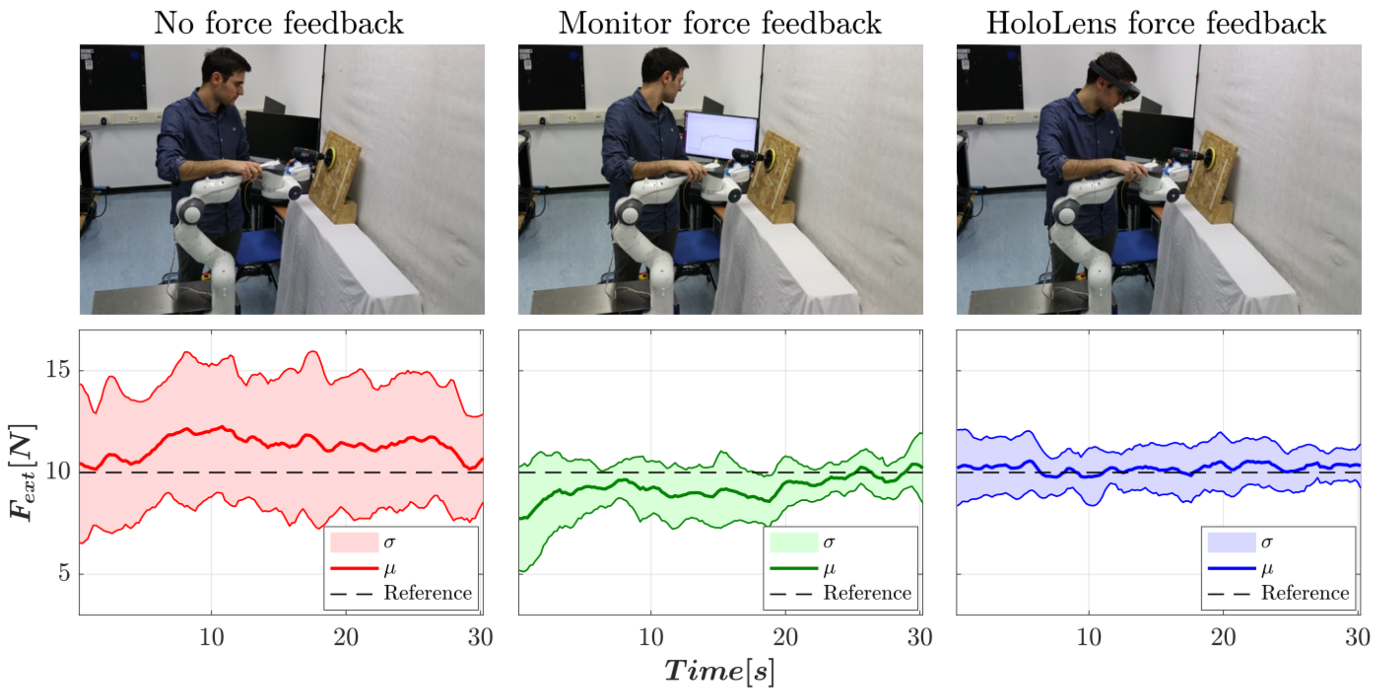

Fig. 5. Experimental collaborative polishing task (IV-B). The photo above show the experimental setup in the three different experimental settings specified by the title. The task force in the three experimental settings is depicted in the plots below. The average value of the force (μ) is shown by the marked line, the standard deviation (σ) by the faded area, and the reference value (10 N) by the dotted line.

far from the desired force, and yellow in between the two) and dimension according to the force value (the higher the force, the longer the arrow) (see Fig. 1). The force feedback corresponds to the estimated force applied to the end-effector of the robot[1]. Participants were asked, after the execution of each trial, to complete a questionnaire based on the Likert scale. The questionnaire was as follows:

Q.1 The polishing was easy to perform;
Q.2 It was difficult to maintain the desired force;
Q.3 It was easy to perceive the amount of force applied;
Q.4 It was difficult to follow the desired trajectory;
Q.5 It was easy to keep the focus on the task execution;
Q.6 It was physically tiresome to accomplish the task;
Q.7 It was psychologically tiresome to accomplish the task;
Q.8 Overall, I felt satisfied with the current task performance.

A score is associated with each possible answer, from strongly disagree to strongly agree, with an assigned score of -5 and +5, respectively. The experiments were approved by the ethics committee Azienda Sanitaria Locale Genovese (ASL) N.3 (Protocollo IIT_HRII_001 (rif. interno:108/2018)).

V. RESULTS

To compare the measured task force between the three different experimental settings, we computed the average and the standard deviation among the ten subjects throughout the experiment for each trial: 1) No force feedback; 2) Monitor force feedback; 3) HoloLens force feedback. The results shown in Fig. 5, suggest that the subjects performed more accurately the polishing task with a feedback interface. In fact, in trials 2 and 3 the standard deviation of the task force is lower with respect to trial 1. However, the average value in trial 2 is still oscillating and it is not centered on the desired value of 10 N. Using the AR interface (trial 3), instead, allows the users to keep an effort almost constant, hence we can conclude that the AR interface provide a better feedback than the monitor. To further support this hypothesis, we performed a one-way analysis of variance (ANOVA test). The boxplots of each trial are displayed in Fig. 6. It resulted that the differences between the data in trial 2 and 3 are statistically significant (p-value < 0.01). These improvements

[1]A video of the experiment is available in https://youtu.be/4MSIy4TC6zs

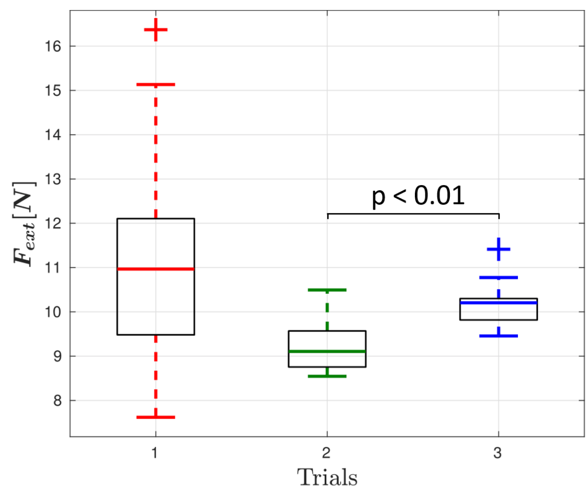

Fig. 6. Experimental collaborative polishing task (IV-B). Boxplots of the task force average value in the three experimental settings (trials). The p-value shows that Trial 3 is statistically significant w.r.t. Trial 2.

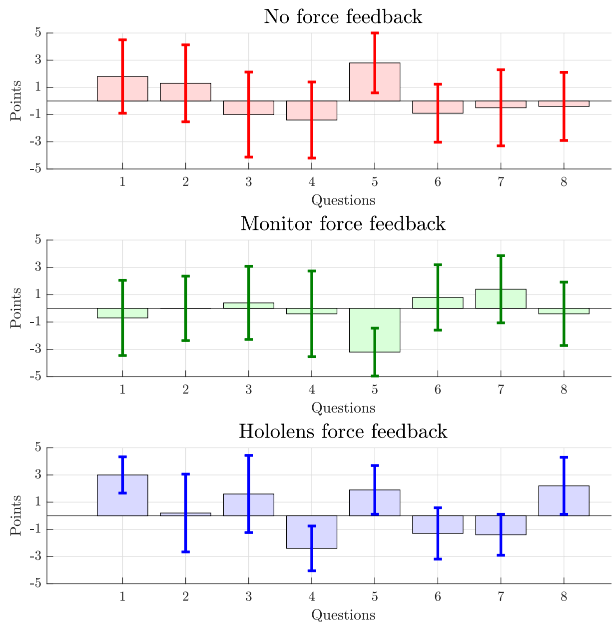

Fig. 7. Likert scale questionnaire scores for the collaborative polishing experiment (IV-B) in the three experimental settings.

are confirmed by the questionnaire results (Fig. 7). In the answers to the question 5 (Q.5) the subjects stated that for them it was easy to keep the focus on the task execution in the third trial, unlike the second trial. The second trial is the most physically tiring for the subjects (Q.6) and it is the least easy to perform (Q.1), but it introduced improvements in the perception of the applied force compared to the first trial (Q.3). However, the subjects confirmed that with HoloLens the perception of the force applied is improved even more. Also with regard to the trajectory followed during the polishing task the subjects prefer AR feedback with respect to the other approaches (Q.4). In particular the trial with force feedback displayed on the monitor was the most difficult to accomplish because the subjects could not focus on the task because of the monitor placement. Overall, the participants felt satisfied with the proposed AR interface (Q.8) and the amount of psychological effort required (Q.7). An interesting suggestion came from one of the participants who advised us to show in the HoloLens display, instead of the current value of force, the force signal as in the case of the monitor (see Fig. 1 and 5).

VI. CONCLUSION

In this paper, we presented an AR interface used in Human-Robot collaborative tasks. The proposed system architecture enables real-time communication between the user and the robot. Our AR interface can introduce many benefits for human operator in HRC settings, in terms of worker performance and situational awareness. Moreover it can be used as an intuitive tool for industrial task scheduling. The results provided in Sec. V supported the effectiveness of our AR interface and suggested that the subjects were satisfied with the task carried out with the AR feedback. Future works will focus on the improvement of the presented interface in different industrial applications, for instance, to coordinate different tasks executed by multiple robots.

ACKNOWLEDGEMENT

This work was supported in part by the Amazon Research Awards 2018.

REFERENCES

[1] 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*, pp. 1–19, 2018.

[2] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, 2018.

[3] A. Cherubini, R. Passama, A. Crosnier, A. Lasnier, and P. Fraisse, "Collaborative manufacturing with physical human–robot interaction," *Robotics and Computer-Integrated Manufacturing*, vol. 40, 2016.

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

[5] A. D. Luca, A. Albu-Schaffer, S. Haddadin, and G. Hirzinger, "Collision detection and safe reaction with the DLR-III lightweight manipulator arm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2006, pp. 1623–1630.

[6] W. Kim, J. Lee, L. Peternel, N. Tsagarakis, and A. Ajoudani, "Anticipatory robot assistance for the prevention of human static joint overloading in human–robot collaboration," *IEEE robotics and automation letters*, vol. 3, no. 1, pp. 68–75, 2018.

[7] L. Peternel, N. Tsagarakis, D. Caldwell, and A. Ajoudani, "Robot adaptation to human physical fatigue in human–robot co-manipulation," *Autonomous Robots*, pp. 1–11, 2018.

[8] S. A. Green, M. Billinghurst, X. Chen, and J. G. Chase, "Human-robot collaboration: A literature review and augmented reality approach in design," *International journal of advanced robotic systems*, vol. 5, 2008.

[9] G. Terenzi and G. Basile, "Smart maintenance–an augmented reality platform for training and field operations in the manufacturing industry," *ARMEDIA Augmented Reality Blog*, 2014.

[10] Boeing tests augmented reality in the factory, https://www.boeing.com/features/2018/01/augmented-reality-01-18.page.

[11] M. Funk, "Augmented reality at the workplace: a context-aware assistive system using in-situ projection," *Dissertation*, 2016.

[12] E. Ruffaldi, F. Brizzi, F. Tecchia, and S. Bacinelli, "Third point of view augmented reality for robot intentions visualization," in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer, 2016, pp. 471–478.

[13] C. Mateo, A. Brunete, E. Gambao, and M. Hernando, "Hammer: An android based application for end-user industrial robot programming," in *2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA)*. IEEE, 2014, pp. 1–6.

[14] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, and A. Ajoudani, "A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant (moca)," *IEEE Robotics and Automation Letters*, 2019, under review.

[15] A. Ajoudani, S. B. Godfrey, M. Bianchi, M. G. Catalano, G. Grioli, N. Tsagarakis, and A. Bicchi, "Exploring teleimpedance and tactile feedback for intuitive control of the pisa/IIT soft hand," *IEEE Transactions on Haptics*, vol. 7, no. 2, pp. 203–215, 2014.

[16] K. B. Reed and M. A. Peshkin, "Physical collaboration of human-human and human-robot teams," *IEEE Transactions on Haptics*, vol. 1, no. 2, pp. 108–120, July 2008.

[17] J. Dumora, F. Geffard, C. Bidard, T. Brouillet, and P. Fraisse, "Experimental study on haptic communication of a human in a shared human-robot collaborative task," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 5137–5144.

[18] K. Baraka, S. Rosenthal, and M. Veloso, "Enhancing human understanding of a mobile robot's state and actions using expressive lights," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2016, pp. 652–657.

[19] E. Lamon, A. De Franco, L. Peternel, and A. Ajoudani, "A capability-aware role allocation approach to industrial assembly tasks," *IEEE Robotics and Automation Letters*, 2019, under review.