Evaluating an Augmented Reality-based Partially Assisted Approach to Remote Assistance in Heterogeneous Robotic Applications

Davide Calandra
Dip. di Automatica e Informatica
Politecnico di Torino
Turin, Italy
davide.calandra@polito.it

Alberto Cannavò
Dip. di Automatica e Informatica
Politecnico di Torino
Turin, Italy
alberto.cannavo@polito.it

Fabrizio Lamberti
Dip. di Automatica e Informatica
Politecnico di Torino
Turin, Italy
fabrizio.lamberti@polito.it

***Abstract*—Among the countless applications of Augmented Reality (AR) in the industry, remote assistance represents one of the most prominent and widely studied use cases. Recently, the way in which assistance can be delivered started to evolve, unleashing the full potential of such technology. New methodologies have been proposed able to foster operators' autonomy and reduce under-utilization of skilled human resources. This paper studies the effectiveness of a recently proposed approach to AR-based remote assistance, referred to as partially assisted, which differs from the traditional step-by-step guidance in the way the AR hints are conveyed by the expert to the operator. The suitability of this approach has been proved already for a number of simple industrial tasks, but a comprehensive study has yet to be performed for validating its effectiveness in complex use cases. This paper addresses this lack by considering as a case study the mastering of a robotic manipulator, a procedure involving a number of heterogeneous operations. The performance of the partially assisted approach is compared with step-by-step guidance based on both objective and subjective metrics. Results showed that the former approach could be particularly effective in reducing the time investment for the expert, allowing the operator to autonomously complete the assigned task in a time comparable to traditional assistance with a negligible need for further support.**

***Index Terms*—augmented reality, industrial applications, remote assistance, autonomous operation, process efficiency**

I. Introduction

Traditional approaches to remote assistance relying on voice/video conferencing greatly limit how the remote expert can guide the on-site operator, especially when 3D spatial referencing or action demonstrations are required [1]. These ones are the situations in which AR potential comes into play.

The use of remote assistance with AR has been largely investigated [2], since it can reduce the need to have experts on-site and, consequently, save time and costs [3].

In such scenarios, the expert is usually required to guide the operator step-by-step till the end of the assistance. This workflow results in a one-to-one mapping between the time invested by the expert and the time required by the operator to complete the task, which may lead to an under-utilization of skilled human resources [4].

To cope with this limitation, a different assistance approach was recently presented [4]. In this approach, referred to as *partially assisted*, the remote expert first delivers all the information required to cope with the given issue in the form of AR contents. Then, the on-site operator executes the provided procedure, without the need for the expert to remain on-line. In this way, the expert is kept busy only for the time that is strictly necessary. According to the definition given in [5], this approach is essentially based on an asynchronous collaboration paradigm, with the possibility to switch to a synchronous paradigm on operator's request.

To make this approach viable, the operator has to be supported by the remote assistance application during the whole procedure, keeping all the AR information provided by the expert accessible over time. For this reason, an AR-powered remote assistance platform was presented along with the approach, and used to evaluate it against the traditional step-by-step, *fully assisted*, guidance.

The partially assisted approach proved to be capable of reducing the duration of the expert intervention in a number of simple and homogeneous industrial tasks, allowing the operator to complete the considered procedures autonomously [4]. However, complex, heterogeneous tasks that could be common in industrial practice were not object of that investigation.

The goal of the current work is to investigate the applicability and performance of the partially assisted approach in a more complex and representative use case, by comparing it with the traditional assistance through a comprehensive user study. The evaluation methodology was inspired by the one used in [4], with the inclusion of new research dimensions. In this case, the mastering procedure of an industrial robot was chosen as a use case. The evaluation was carried out in both subjective and objective terms, considering aspects like the time saved by the remote expert, the time required to complete the whole procedure (beyond the end of the call), as well as many factors pertaining the user experience.

II. Related works

Today, among the numerous application fields of AR in industry, one that is attracting a lot of attention by both researchers and companies is remote assistance [6].

A first example of works supporting remote assistance with AR is represented by [7], in which two wearable systems based on an HMD and a chest-worn display were developed to support the collaboration between the operator and the remote expert. The work mainly focused on the advantages brought by the use of an AR-based *laser pointer*, a visual tool that allows the expert to point a specific target while assisting the operator.

Besides the pointer, other AR-based tools have been introduced to support remote assistance. As a matter of example, the work in [8] proposed a system for task guidance based on different types of *annotations*, like 3D models, shapes (i.e., arrows, boxes, and circles) and hand drawings. In order to enhance the clarity of voice instructions, the expert has the possibility to capture some frames of the video transmitted in real time by the operator and add annotations to it.

In [9], another system targeted to industrial maintenance facilitates remote cooperation and training by allowing the expert to place AR annotations by using a marker-based approach. Annotations mainly consist of *anchored texts* that are used to supplement voice instructions, as they convey many information that can be also exploited after the assistance.

Although with markers it is possible to obtain robust and precise tracking that can be very effective for many different industrial applications [10], this technology suffers from some limitations when applied to remote assistance [11]. For this reason, alternative marker-less technologies for anchoring AR contents were also investigated in the literature. For instance, it is possible to take into account the Simultaneous Localization and Mapping (SLAM). This technology leverages so-called feature points computed on the video feed to estimate the position of an AR device in the real world [12], enabling the superimposition of digital assets on physical objects.

Nowadays, the wide availability of sensors and the computational power of common smartphones and tablets make these devices suited to support AR applications based on SLAM, lowering down the cost of hardware. As a matter of example, the system in [13] makes use of SLAM to attach visual assets, such as text and drawings, on the real objects. The system allows the remote expert to observe the real-time, first-person view of the operator's site. Annotations created by the expert can be visualized by the operator on Microsoft's HoloLens.

Although all the systems reviewed above allow remote experts to successfully provide assistance by means of different AR tools, most of them lack the possibility to retain a chronological list of the received instructions/guidance (e.g., in the form of a timeline). The idea of introducing an ordered list of received instructions was explored in [14]. The proposed remote assistance platform provides the operator with a navigable timeline of annotated snapshots, which can be referenced during the whole session.

For what it concerns contents reuse, in [15], an AR framework for machinery inspection was presented. The framework is capable to record data about the intervention while the expert is providing instructions to the operator. The data that remain available to the operator after the end of the session can be photos, texts or audio clips. The back-end of the framework allows the expert to reconfigure the support for different scenarios, by letting him or her directly upload manuals and instructions to be used for further assistance.

Based on the analysis of scientific works, the authors [4] realized that the AR features presented above could be leveraged to support a shift from step-by-step assistance to a more autonomy-oriented approach, named partially assisted. Hence, they developed a single platform offering all the analyzed AR features and services needed to support this methodology. The devised platform relies on AR-enhanced audio-video calls between experts and operators. From a dedicated web portal, the expert can add to the received video feed a number of temporary or persistent AR contents, either anchored or overlaid to it, to improve the richness/clearness of the instructions. Augmented contents are visualized by the operator on a mobile device that leverages the SLAM technology to anchor contents on real objects. The platform integrates a chronological timeline of the delivered/received instructions and offers the possibility to retain access to AR contents after closing the call as well as to re-establish it without leaving the AR scene. The expert can use the portal to schedule/manage the assistance and prepare the instructions in advance. Finally, the mobile application allows the operator to record both the audio-video stream and the instructions timeline for future consultation.

In [4], the effectiveness of the proposed approach was assessed by comparing it with step-by-step guidance in three simple use cases involving a collaborative robot, namely, the assembly of a gripper (GA), the execution of a load data determination (LD) procedure, and the recovery from an emergency stop (ER). The selection of three simple, homogeneous tasks made it possible to isolate the specific effects that the assistance approach can have on each of them. In particular, for the GA task, characterized by familiar and repetitive actions, the partially assisted approach showed the best performance when compared with the traditional guidance, reducing the time invested by the expert by $\sim$62%. The ER task, characterized by complex and unfamiliar actions, showed the lowest advantage ($\sim$20% time reduction), at the cost of a slightly longer procedure completion time. Results for LD showed intermediate performance ($\sim$36% time reduction). The evaluation also considered several subjective dimensions, such as perceived user performance, system performance, learnability, memorability, frustration and appropriateness. In that case, clear advantages of the partially assisted approach were only observed for the GA task and for the overall user-assigned scores (averaged among the three tasks).

The previous evaluation did not provide any information about the suitability of the partially assisted approach to heterogeneous remote assistance use cases. In fact, although

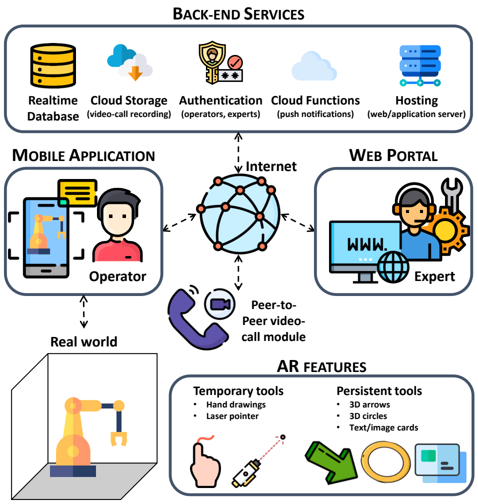

Fig. 1. Architecture of the remote assistance platform.

(a) Hand drawings (b) Laser pointer (c) 3D arrows (d) 3D circles

(e) Text cards (f) Image/GIF cards (g) Instruct. timeline (h) Record./Playback

Fig. 2. AR tools available in the mobile application.

the tasks considered in [4] were common industrial procedures, they do not fully represent the complexity of a generic remote assistance scenario. For this reason, the experimental evaluation carried out in this paper considers a complex task, i.e., the mastering of an industrial robot, encompassing different types of operations and underlying concepts, with the final aim of assessing to what extent previous results could generalize.

III. Remote Assistance Platform

As said, the analysis methodology used in this work was inspired to the one adopted in [4]. The same AR-powered remote assistance platform was used as a common ground for the evaluation (its architecture is shown in Fig. 1). The system includes three main components: the operator-side mobile application, the expert web portal, and the back-end services. In the following, each module is briefly described.

A. Mobile Application

The operator-side mobile application runs on any Android mobile device, e.g., tablets and smartphones, with an internet connection. The AR functionalities were developed with the AR Core library[1]. The remote assistance takes place via an audio-video call. The camera of the mobile device handed by the operator is used to provide the expert with the video feed on which AR-based tools can be anchored.

The platform supports the most common AR-based tools reviewed in the state of the art, for which some usage examples are shown in Fig. 2. The tools can be categorized into two different groups: temporary and persistent. To the first group belong those tools that are displayed as 2D overlays over the video feed, i.e., the hand drawings (Fig. 2a) and the laser pointer (Fig. 2b). Temporary means that these elements vanish a few seconds after placement. The group of persistent tools includes 3D shapes, such as arrows (Fig. 2c) and circles (Fig. 2d), as well as instruction cards (Figs. 2e and 2f) that can be placed over real-world elements by leveraging 6-DOF tracking. Instruction cards can contain text or images and can be displayed as 2D overlays or spatially anchored. Once inserted, cards can be visualized into a scrollable timeline (Fig. 2g). The recording of the audio-video call is saved and can be visualized after the assistance.

B. Web Portal

The interface of the web portal for providing the remote assistance includes several components, such as the context data concerning the operator and the assistance, the video feed of the operator camera, the tool palette for inserting the desired AR tools, a panel to create new instruction cards on-the-fly, the list of previously created instruction cards, and the instruction timeline of the current session. The audio and video communication between the operator and the expert is established via the Twilio APIs[2].

C. Back-End Services

A number of back-end functionalities are made available to both the expert and the operator using off-the-shelf features of Google Firebase[3]. In this way, the developed networked platform supports user registration and authentication, the scheduling of audio-video calls, the hosting of session recording, the use of OCR as well as of push notifications.

IV. Experimental Setup

As mentioned, the goal of this work is to evaluate the effectiveness of the partially assisted approach for the execution of a complex task characterized by heterogeneous operations and concepts. To this aim, a new user study inspired by the experimental evaluation presented in [4] was performed by involving 23 volunteers. The same experimental protocol described in [4] was adopted. However, a more complex task requesting participants to deal with more heterogeneous activities within the same procedure was considered, and

[1]Google ARCore: https://developers.google.com/ar

[2]Twilio APIs: https://www.twilio.com/

[3]Google Firebase: https://firebase.google.com/

new metrics were collected to explore additional research dimensions. The partially assisted approach was compared with the step-by-step guidance by considering both objective and subjective aspects, as detailed in the following.

A. Participants

The 23 volunteers (10 male and 13 female) were aged between 24 and 51 ($\mu = 33.52$, $\sigma = 8.16$). The majority of them (16 subjects) were students and academic staff at the authors' university; the remaining (7 subjects) were recruited among the administrative staff of KUKA Roboter Italia Spa. Information collected through a demographic questionnaire revealed that participants were familiar with the Android environment and had a medium knowledge of audio-video conferencing applications. Moreover, data showed that selected subjects had a very low previous experience with AR applications and tools (both software and hardware) for controlling industrial robots. The limited skills in the field of robotics made participants suitable for playing the role of novice users.

B. Methodology

As anticipated, participants were asked to complete a demographic questionnaire aimed at evaluating their familiarity with the technologies used in the experiment. Then, details regarding the mobile application, the industrial robot and the SmartPad (the device used to manually control the robot) were presented to the participants. They were left free to use the SmartPad, the smartphone used to execute the mobile application (a Huawei HONOR 8X), and the application itself. Finally, the experiment goals were presented.

In particular, participants were requested to play the role of a generic operator requesting support for a given procedure on the robot (more details about the procedure are provided in Section IV-C). The role of the expert was covered by a technician in charge of customer service at KUKA. Once the application was launched on the mobile device and data regarding the robot were collected through the troubleshooting phase, participants were requested to start the audio-video call to receive the remote assistance.

As said, the remote expert provided assistance in two different modalities. In the fully assisted modality, later abbreviated *FA*, the operator received continuous support from the expert until the completion of the procedure. That is, the instructions were provided step-by-step by leveraging the voice/video conferencing and AR-based tools.

With the partially assisted modality, in the following abbreviated *PA*, the expert provided all the required instructions at the beginning of the call, then asked the operator to proceed autonomously. To complete the procedure, the operator had to largely utilize the instructions sent by the expert. In the cases in which the issues cannot be solved only with the received instructions, the operator had the possibility to call the expert again in order to receive further hints. No threshold was set during the experiments on the number of calls or on the time before recalling. Nevertheless, the impact of these kinds of restrictions could be worth investigation in the future.

Participants were split into two groups: half of them were requested to complete the task in the FA modality, the other half in the PA modality. Assignment to groups was made by balancing as much as possible the distribution in terms of age, gender, and previous experience.

During the experiment, objective data regarding the participants' performance were logged. Finally, after the completion of the procedure, participants were requested to fill in a post-test questionnaire in order to evaluate their experience in subjective terms. More details about the objective and subjective metrics are discussed in Section IV-D.

C. Use Case

In order to compare participants' performance and evaluate their experience on a common ground characterized by a representative level of complexity, the procedure for the *mastering* of a KUKA's KR 16 industrial robot was selected.

The goal of the procedure is to synchronize the mechanical zero and the electrical zero positions for each axis of the robot [16]. This procedure represents a fundamental step to make the robot operate in optimal conditions, since it allows the robot to accurately follow the programmed trajectories.

The mastering procedure it is quite representative of a wide range of operations that could have to be performed in a robotic (and industrial) assistance, since it requires to move some equipment and inspect it, manipulate other devices, perform some configurations by interacting with a user interface, etc. Its key steps can be summarized as reported below.

1) Each joint of the robot has to be moved to the pre-calibration position in order to bring each axis approximately close to its mechanical zero. This position can be reached by visually aligning, for each axis, two specific visual indicators located near the axis joints that are named "notches" (an example of aligned notches is shown in Fig. 3a). In order to make the robot attain the pre-calibration position, the user operates on the corresponding $+$ and $-$ buttons of the SmartPad.
2) Then, the user has to tighten an external portable sensor named Electronic Mastering Device (EMD) on the corresponding measuring cartridge of the axis to be calibrated. An example showing the EMD attached to the cartridge of an axis is illustrated in Fig. 3c.
3) Once the EMD has been attached, the user electrically connects it to the robot.
4) The user configures a semi-automatic program to calibrate the axis the EMD is connected to interacting with the SmartPad touchscreen (Fig. 3b).
5) Once the program has been configured, the user activates it by pressing at the same time the Consent button and the Start button (Fig. 3b) and keeping them pressed while the axis moves to be calibrated.
6) The user disconnects the cables from the EMD and removes it from the measuring cartridge.
7) The user repeats steps 2–6 for all the remaining axes to be calibrated. It is possible to maintain the connection

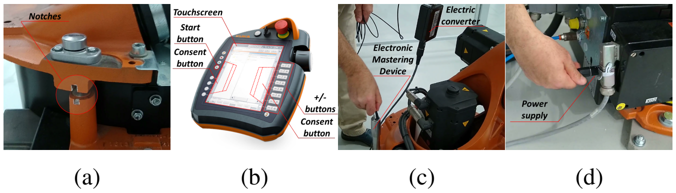

Fig. 3. Mastering procedure for calibrating the KR 16 robot.

between the power supply and the converter, since only the EMD needs to be moved.

Since step 7 is basically a repetition of the previous operations for all the robot axes, it was decided to limit the procedure to steps 1–6 for the experiments in order to enable a fair comparison of the FA and PA modalities.

In the FA modality, the operator is assisted till the end of the procedure (until step 6). In the PA modality, the expert supports the operator in the execution of step 1, then provides all the instructions required to complete the remaining steps (2–6) inviting the operator to familiarize with them; afterwards, he or she asks the operator to complete the procedure autonomously.

A single expert was employed for all the experiments. The two assistance modalities made the expert use different tools for supporting the operators. In the FA modality, the expert was able to provide all the needed information by using only temporary, explanation-complementary tools, such as the laser pointer (65.39% of times), hand drawings (19.23%), and 3D shapes (15.38%); instruction cards were not used since the operations to perform were explained verbally. In the PA modality, most of the support was provided by exploiting persistent, self-explanatory instruction cards (77.42% of times), rarely the temporary tools (12.91% the laser pointer, 9.67% the hand drawings), and never the 3D shapes. Text/image cards used for the assisted part of the PA modality were identical for all the participants, since they were created in advance by the expert in the web portal.

Two videos showing an example of assistance with the two modalities are available for download[4].

D. Evaluation Criteria

In order to evaluate participants' performance and experience, both objective and subjective metrics were collected.

Regarding objective evaluation, as in [4], two metrics were considered. The first one, named *call time*, measures the overall duration of the audio-video call between the operator and the expert. The second metric, named *completion time*, takes into account the time spent by the operator to execute all the steps of the procedure. Due to the fact that in the FA modality the actions carried out by the operator were supervised till the end of the task by the expert, the completion time corresponds to the call duration metric. In the PA modality, the call duration metric accounts for the possible additional calls requested by the operators, as it represents the cumulative duration of the assistance. The number of re-calls was also recorded.

For what it concerns the subjective evaluation, it was performed by requesting participants to fill in a post-test questionnaire (available for download[5]). Differently than in [4], the first two sections of the questionnaire were aimed to evaluate the overall usability of the application in the two modalities. The first section included questions from the standard System Usability Scale (SUS) [17], whereas the second, based on the Post-Study System Usability Questionnaire (PSSUQ) [18], was aimed to evaluate users' perceived satisfaction about System Usefulness (SYSUSE), Information Quality (INFOQUAL) and Interface Quality (INFQUAL). The third section, which included specific statements to assess perceived user performance, system performance (in terms of quality of the audio-video communication), learnability, memorability, and frustration, roughly corresponded to questionnaire already used in [4], with some modifications. In particular, the questions related to the appropriateness were not included, since in [4] they were used to directly compare multiple tasks, whereas in this work only one task is considered. Finally, few redundant questions related to some of the previously mentioned dimensions were discarded too.

In order to be compliant with the SUS and PSSUQ questionnaires, statements in the first and second section had to be rated, respectively, on a 5-point Likert scale and a 7-point Likert scale. For the last section, a 5-point Likert scale was again adopted. For the sake of simplicity, the three scales were standardized to make them go from total disagreement (1) to total agreement (5 or 7). Space for comments was provided.

V. Results

The objective and subjective results are discussed below.

A. Objective Results

Results regarding completion time and call duration are illustrated in Fig. 4. Unpaired samples t-test with 5% significance ($p < 0.05$) was used to analyze significant differences.

Concerning the call duration, Fig. 4 shows that the PA modality made the subjects more confident that they could complete the procedure autonomously. On average, for the PA group the time invested by the expert in the assistance was 31.50% shorter than for the FA group ($p = 0.0001$). In the PA modality, only four subjects needed further help to complete the task after the first call was terminated; three of them reactivated the call once, whereas one of them needed two additional calls to complete the task. These results confirmed the suitability of the proposed approach for this kind of tasks, since only a limited number of subjects needed to re-call to finish the procedure. Regarding this metric, observed values are in line with those reported in [4] for tasks characterized by high levels of complexity and unfamiliarity.

For what it concerns completion time, no statistically significant difference was found between the two group. This outcome can be regarded as a confirmation of the fact that the paradigm shift introduced by the studied approach did not impact on the overall operators' performance.

[4]Videos of the two modalities: http://tiny.cc/7xwutz

[5]Questionnaire: http://tiny.cc/6xwutz

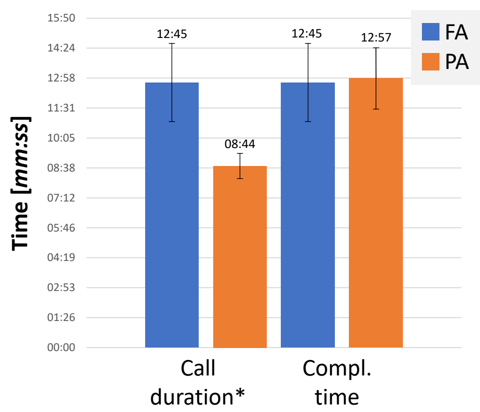

Fig. 4. Objective results concerning call duration and average completion time (standard deviation expressed through error bars). Significant results are marked with a * symbol.

B. Subjective Results

Statistical significance was tested with the same methodology adopted for the objective evaluation.

Starting with results for SUS [17], subjects perceived both the modalities as characterized by a high usability (77.72 for FA vs 78.75 for PA, $p = 0.1847$), with scores that were greater than 71.1 (threshold for *Good*). No statistically significant difference was found between the two modalities: based on the categorization in [19], both the FA and the PA modalities were rated as grade B+.

For the PSSUQ, results obtained by averaging the scores for the different sub-scales only showed a statistically significant difference for the SYSUSE (6.65 for FA vs 6.89 for PA, $p = 0.0275$). Possible motivations for this result can be identified by analyzing more in detail the average scores assigned to individual statements in Table I.

Considering the SYSUSE sub-scale, it can be observed that subjects perceived to be faster with the PA modality than with the FA modality (statement 3, 5.27 vs 7.00, $p = 0.0001$). This finding, which confirms objective results, is probably due to the fact that, with the FA modality, subjects were requested to continuously interrupt their operations in order to listen/receive the step-by-step instructions provided by the remote expert. When operating with the PA modality, they had the possibility to receive all the needed information and then to execute the required operations without breaking the flow of their actions. Significant differences were found also for statements 12 and 15 belonging to the INFOQUAL and INTQUAL sub-scales. In particular, for the former statement, subjects found that the organization of the AR elements was clearer in the PA modality than in the FA modality (5.45 vs 6.42, $p = 0.0338$). This result could be related to the fact that most of the AR instructions provided with the PA modality (images and texts) were also inserted in the timeline for chronological consultation, whereas the AR tools used in the FA modality were not recorded in it. The presence at the same time of multiple AR hints could create confusion, especially when the instructions are provided with groups of very close and partially overlapping elements, which the user can only hide or show in bulk. As for the latter statement on the presence of all the expected functionalities, the subjects preferred the PA modality to the FA modality (6.27 vs 6.92, $p = 0.0290$). This result may be due to the fact that, in the FA modality, most of the instructions were provided with voice, and subjects could have expected a more prominent role of AR in the assistance, as well as more differences with respect to a common audio-video call.

Results of the last section of the questionnaire are tabulated in Table II. Focusing on statistically significant results and considering first the user performance dimension, it can be observed that, based on statement 2, subjects reported that they needed less support when using the PA modality than the FA modality (4.63 vs 5.00, $p = 0.0207$). As suggested in [4], this result could be actually related to the way instructions were delivered in the two modalities. In the FA modality, where they were delivered step-by-step, it happened that subjects asked the expert for information yet to be provided (e.g., because included in a following step). In the PA modality, all the required instructions were provided at the beginning; hence, during the execution of the task the subjects already knew every detail of the whole procedure; these contents were also retained within the instruction timeline, making subjects feel that no further help from the expert was needed.

Concerning learnability (here intended as the perceived ability of the system to transfer knowledge about the task to be performed), participants judged the PA modality as better than the FA in transferring both theoretical concepts (statement 4, 3.45 vs 5.00, $p = 0.0001$) and practical operations (statement 5, 4.09 vs 5.00, $p = 0.0002$). For this reason, these contents were perceived as easier to learn. Subjects guided step-by-step with FA, did not actually required to memorize or understand what they were doing. On the other hand, with the PA modality, participants were requested to operate autonomously, therefore they needed to memorize and understand all the required operations. This outcome is also confirmed by the scores assigned to statement 6, which indicate that participants felt to be following the instruction more "mechanically" (i.e., without understanding the reasons behind them) with FA with respect to PA (2.45 vs 1.16, $p = 0.0012$).

Besides the improved capability of the PA modality to make information about the procedure easier to learn and understand, results concerning memorability (intended as the ability to remember information needed to perform the task again) indicated that this modality also made participants felt to be able to re-execute the assigned task using only the information remaining on the device after the call (instructions timeline and audio-video call recording) both right after completing it (statement 7, 3.63 vs 5.00, $p = 0.0001$) and in the future (statement 8, 2.73 vs 4.67, $p = 0.0001$). The reason may be related to the fact that the PA modality allowed participants to partially experience the possibility to work autonomously.

Finally, it can be observed that participants found that making the remote expert wait until they had completed the operations related to the current instruction may be a factor

TABLE I: Subjective results about post-study system usability based on the PSSUQ [18]. The higher the score, the higher the agreement (original PSSUQ scale reversed for sake of clarity). Statistically significant differences ($p < 0.05$) are marked with a * symbol, whereas bold font indicates the best of the two values

Sub-scale	Statement	FA	PA	*p-value*
SYSUSE	1. Overall, I am satisfied with how easy it is to use this system	6.73	6.75	0.9323
	2. It was simple to use this system	7.00	6.75	0.3501
	3. I was able to complete the task quickly using this system*	5.27	**7.00**	0.0001
	4. I felt comfortable using this system	7.00	7.00	NA
	5. It was easy to learn to use this system	6.91	7.00	0.3071
	6. I believe I could become productive quickly using this system	7.00	6.83	0.3502
INFOQUAL	7. The system gave information that clearly told me how to fix problems	7.00	7.00	NA
	8. Whenever I made a mistake using the system, I could recover easily and quickly	6.91	6.92	0.9515
	9. The information (Troubleshooting FAQs) provided with this system was clear	7.00	6.83	0.3502
	10. It was easy to find the information I needed	6.91	6.83	0.7013
	11. The information was effective in helping me complete the task	6.73	6.75	0.9323
	12. The organization of (AR) information on the device screens was clear*	5.45	**6.42**	0.0338
INTERQUAL	13. The interface (device, interaction means, etc.) of this system was pleasant	6.45	6.42	0.9041
	14. I liked using the interface (device, interaction means, etc.) of this system	4.45	4.25	0.7952
	15. This system has all the functions and capabilities I expect it to have*	6.27	**6.92**	0.0290
	16. Overall, I am satisfied with this system	6.82	6.92	0.6176

TABLE II: Subjective results about the five dimensions included in the custom section of the questionnaire. The higher the score, the higher the agreement. Statistically significant differences ($p < 0.05$) are marked with the * symbol, whereas bold font indicates the best of the two values

Category	Statement	FA	PA	*p-value*
User performance	1. I think that without the AR hints my performance would have been worse	4.91	4.92	0.9515
	2. I did not need any further help from the expert when performing the given task*	4.63	**5.00**	0.0207
System performance	3. I thought that the audio-video communication quality was adequate	4.90	4.83	0.6098
Learnability	4. The received assistance let me effectively learn required concepts*	3.45	**5.00**	0.0001
	5. The received assistance let me effectively learn procedural step*	4.09	**5.00**	0.0002
	6. I performed the operations in a "mechanical" way, without understanding the concepts/reasons*	2.45	**1.16**	0.0012
Memorability	7. I think that I would be able to repeat the procedure alone right now, using the information available on the device (session history)*	3.63	**5.00**	0.0001
	8. I think that I would be able to repeat the procedure alone in the future, using the information available on the device (session history)*	2.73	**4.67**	0.0001
Frustration	9. The kind of received assistance put pressure on me*	2.42	**1.27**	0.0023

that could create frustration in executing the procedure. In fact, according to statement 9, subjects stated that the received assistance put more pressure on them with the FA modality than with the PA modality (2.42 vs 1.27, $p = 0.0023$).

VI. Discussion and Conclusion

This paper evaluated a recently proposed approach to remote assistance in a complex and heterogeneous industrial use case. This approach, named partially assisted (PA), aims at reducing the time invested by an expert while providing support to on-field operators through AR-powered remote assistance tools. This is done by providing and explaining all the required instructions at the beginning of the assistance session, and then letting the operator execute the operations autonomously till the completion of the illustrated steps.

Asynchronous collaboration paradigms have been rarely studied in the literature, as opposed to synchronous paradigms [5]. Hence, further investigations aimed to compare them were needed. To this aim, the PA approach supports both the paradigms, even though it is mainly oriented to the asynchronous one. In [4], the authors showed the applicability and advantages of PA over the synchronous, fully assisted (FA) approach when dealing with simple industrial tasks. In particular, the performance of the proposed approach appeared to be proportional to the familiarity and repetitiveness of the procedure, and inversely proportional to its complexity.

In this work, a comparable evaluation strategy is adopted to evaluate the PA approach on the mastering of an industrial robot. This task encompasses a wider range of different concepts and operations, basically representing a blend of features found in the previously considered tasks.

Results confirmed that the PA approach can significantly reduce the time of the remote expert intervention also in this kind of tasks, letting the operators complete the procedure autonomously without any significant increase in the total operation time. Moreover, operators who used the PA approach perceived the system as significantly more useful, capable to make them work more quickly, efficiently, and relaxedly, and to better convey the expert's knowledge. These outcomes are in line with the results reported in [4] regarding tasks characterized by high familiarity and repetitiveness. This finding suggests that, for heterogeneous tasks including both familiar and unfamiliar concepts as well as simple and complex operations, the benefits brought by the PA approach under favourable conditions (simple and familiar steps) are much more prominent than the disadvantages on unfavourable ones (complex and unfamiliar steps).

The selected use case included a repetitive component, i.e., the mastering of each of the six robot axes, which was removed for the experiment. This choice was made to consider the worst-case scenario. Hence, it is reasonable to assume that the significant difference between the two approaches could be even more pronounced when executing the whole procedure (which is the one that is commonly performed on this kind of robots). Moreover, a large number of industrial procedures are characterized by repetitions, or by sequences of similar operations that do not require separate explanations. Therefore, the proposed approach, if well applied and correctly supported by the described AR tools, can likely bring the observed benefits to many of the possible industrial use cases.

It is worth observing that, if on the one hand, the adoption of the PA approach reduces the time spent by those providing remote support to deliver the support, on the other hand it also changes the complexity of their work. This aspect could introduce new factors to be considered by the companies, such as the increased cost of the support, the new skill requirements for the involved subjects, etc.

The main limitation of the current evaluation is the adoption of only one procedure to illustrate the suitability of the PA approach for complex use cases. The selected procedure may not be representative of any possible situation. Hence, in other scenarios it may be more difficult for experts to predict all the problems, and the number of re-calls may be larger. Moreover, like in the previous work, it is assumed that the expert is already in possession of all the AR contents to be provided in the upcoming assistance session.

Future developments will be aimed at evaluating the PA approach for other complex use cases, as well as in less favourable conditions. For example, in case that part of the required instructions have to be generated on-the-fly by the expert to react to unexpected conditions, thus better representing real-world scenarios.

Furthermore, aspects such the short and long term memorability have been only assessed in a subjective way, by asking participants if they thought they would have been able to execute the procedure right after its completion or in the future. To evaluate them in objective terms, a new study may be reiterated in the future, to confirm the participant perceptions by asking them to repeat the same procedure autonomously.

Finally, a further investigation may be performed on the expert-side of the remote assistance platform. By involving a statistically representative sample of remote operators, the PA may be evaluated from the experts' viewpoint, thus providing a more general characterization of the approach with respect to the company staff actually offering the assistance.

Acknowledgment

This work has been developed in the frame of the VR@POLITO initiative. The authors want to acknowledge the help given by KUKA Roboter Italia Spa which made available the equipment used in the experiments, and of Mauro Cerrato, who contributed to the development of the platform and to the evaluation.

References

[1] C. Elvezio, M. Sukan, O. Oda, S. Feiner, and B. Tversky, "Remote collaboration in ar and vr using virtual replicas," in *ACM SIGGRAPH 2017 VR Village*. ACM, 2017, pp. 1–2.

[2] D. Mourtzis, V. Zogopoulos, and K. Vlachou, "Augmented reality application to support remote maintenance as a service in the robotics industry," *Procedia CIRP*, vol. 63, pp. 46–51, 2017.

[3] L. F. de Souza Cardoso, F. C. Martins Queiroz Mariano, and E. R. Zorzal, "A survey of industrial augmented reality," *Comp. & Industrial Eng.*, vol. 139, p. 106159, 2020.

[4] D. Calandra, A. Cannavò, and F. Lamberti, "Improving ar-powered remote assistance: a new approach aimed to foster operator's autonomy and optimize the use of skilled resources," *The International Journal of Advanced Manufacturing Technology*, vol. 114, no. 9, pp. 3147–3164, Jun 2021. [Online]. Available: https://doi.org/10.1007/s00170-021-06871-4

[5] A. Irlitti, R. T. Smith, S. Von Itzstein, M. Billinghurst, and B. H. Thomas, "Challenges for asynchronous collaboration in augmented reality," in *IEEE International Symposium on Mixed and Augmented Reality*, 2016, pp. 31–35.

[6] A. Gallala, B. Hichri, and P. Plapper, "Survey: The evolution of the usage of augmented reality in industry 4.0," *IOP Conference Series: Materials Science and Engineering*, vol. 521, pp. 1–11, 2019.

[7] N. Sakata, T. Kurata, and H. Kuzuoka, "Visual assist with a laser pointer and wearable display for remote collaboration," *Transactions of the Virtual Reality Society of Japan*, vol. 11, pp. 561–568, 2006.

[8] M. Rice, K.-T. Ma, H. H. Tay, J. Kaliappan, W. L. Koh, W. P. Tan, and J. Ng, "Evaluating an augmented remote assistance platform to support industrial applications," in *IEEE World Forum on Internet of Things*, 2018, pp. 592–597.

[9] T. Alexander, A. Ripkens, M. Westhoven, M. Kleiber, and C. Pfendler, "Virtual tele-cooperation: Applying AR and VR for cooperative tele-maintenance and advanced distance learning," in *Advances in Human Factors in Training, Education, and Learning Sciences*, 2017, pp. 234–244.

[10] E. Battegazzorre, D. Calandra, F. Strada, A. G. Bottino, and F. Lamberti, "Evaluating the suitability of several AR devices and tools for industrial applications," in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, 2020, pp. 1–20.

[11] F. Ferrise, G. Caruso, and M. Bordegoni, "Multimodal training and tele-assistance systems for the maintenance of industrial products," *Virtual and Physical Prototyping*, vol. 8, pp. 113–126, 2013.

[12] C.-W. Chen, W.-Z. Chen, J.-W. Peng, B.-X. Cheng, T.-Y. Pan, and H.-C. Kuo, "A real-time markerless augmented reality framework based on SLAM technique," in *IEEE International Symposium on Pervasive Systems, Algorithms and Networks & IEEE International Conference on Frontier of Computer Science and Technology & IEEE International Symposium of Creative Computing*, 2017, pp. 127–132.

[13] D. Fang, H. Xu, X. Yang, and M. Bian, "An augmented reality-based method for remote collaborative real-time assistance: From a system perspective," *Mobile Networks and Applications*, vol. 25, pp. 1–14, 2019.

[14] M. Rice, S. Chia, H. Tay, M. Wan, L. Li, J. Ng, and H. Lim, "Exploring the use of visual annotations in a remote assistance platform," in *ACM CHI Conference on Human Factors in Computing Systems*, 2016, pp. 1295–1300.

[15] R. Perla, G. Gupta, R. Hebbalaguppe, and E. Hassan, "InspectAR: An augmented reality inspection framework for industry," in *IEEE International Symposium on Mixed and Augmented Reality*, 2016, pp. 355–356.

[16] P. Mako, "Principles of mastering at KUKA robots," *International Scientifica Journal about Machatronics*, vol. 2, pp. 7–12, 2017.

[17] J. Brooke *et al.*, "SUS - A quick and dirty usability scale," *Usability Evaluation in Industry*, vol. 189, pp. 4–7, 1996.

[18] J. R. Lewis, "Usability testing," *Handbook of Human Factors and Ergonomics*, vol. 12, pp. 1267–1312, 2006.

[19] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," *J. of Usability St.*, vol. 4, pp. 114–123, 2009.