

Improving AR-powered remote assistance: A new approach aimed to foster operator's autonomy and optimize the use of skilled resources

*Original*

Improving AR-powered remote assistance: A new approach aimed to foster operator's autonomy and optimize the use of skilled resources / Calandra, Davide; Cannavò, Alberto; Lamberti, Fabrizio. - In: THE INTERNATIONAL JOURNAL OF ADVANCED MANUFACTURING TECHNOLOGY. - ISSN 1433-3015. - STAMPA. - 114:(2021), pp. 3147-3164. [10.1007/s00170-021-06871-4]

*Availability:*

This version is available at: 11583/2873218 since: 2021-06-02T09:43:45Z

*Publisher:*

Springer

*Published*

DOI:10.1007/s00170-021-06871-4

*Terms of use:*

openAccess

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)



# Improving AR-powered remote assistance: a new approach aimed to foster operator's autonomy and optimize the use of skilled resources

Davide Calandra<sup>1</sup> · Alberto Cannavò<sup>1</sup> · Fabrizio Lamberti<sup>1</sup>

Received: 11 October 2020 / Accepted: 5 March 2021 / Published online: 21 April 2021  
© The Author(s) 2021

## Abstract

Augmented reality (AR) has a number of applications in industry, but remote assistance represents one of the most prominent and widely studied use cases. Notwithstanding, although the set of functionalities supporting the communication between remote experts and on-site operators grew over time, the way in which remote assistance is delivered has not evolved yet to unleash the full potential of AR technology. The expert typically guides the operator step-by-step, and basically uses AR-based hints to visually support voice instructions. With this approach, skilled human resources may go under-utilized, as the time an expert invests in the assistance corresponds to the time needed by the operator to execute the requested operations. The goal of this work is to introduce a new approach to remote assistance that takes advantage of AR functionalities separately proposed in academic works and commercial products to re-organize the guidance workflow, with the aim to increase the operator's autonomy and, thus, optimize the use of expert's time. An AR-powered remote assistance platform able to support the devised approach is also presented. By means of a user study, this approach was compared to traditional step-by-step guidance, with the aim to estimate what is the potential of AR that is still unexploited. Results showed that with the new approach it is possible to reduce the time investment for the expert, allowing the operator to autonomously complete the assigned tasks in a time comparable to step-by-step guidance with a negligible need for further support.

**Keywords** Augmented reality · Industrial applications · Remote assistance · Autonomous operation · Process efficiency

## 1 Introduction

Digital technologies are transforming more and more the world economy and the global industry. As a result, concepts such as Industry 4.0 and smart manufacturing have now become commonplace. Technologies like virtual reality (VR) and augmented reality (AR) are playing a big part in this ongoing industrial revolution. AR, in particular, is

showing a growing trend for what it concerns both research interest and patent filing [12, 13, 27, 44].

Many literature works investigated the benefits brought by AR to the industry for various use cases like, e.g., training [8], collaborative design [50], repair and maintenance [3, 4, 55], assembly [9, 31, 42, 51], customer service [18], process simulation and monitoring [10, 26], logistics [25], and quality control [32], proposing and evaluating a number of widely heterogeneous solutions.

Notwithstanding, questions on key factors pertaining the use of AR in industry such as usability, efficiency, effectiveness, and user experience, among others, are still open [5, 56].

As reported in [44], most of the industrial AR applications take advantage of visual cues to show the user how to execute a given task, e.g., in a maintenance or learning scenario. In principle, this approach can be used to replace or supplement paper-based instructions and manuals [9]. However, AR applications cannot completely substitute skills and know-how of experienced employees (at least, not

---

✉ Fabrizio Lamberti  
fabrizio.lamberti@polito.it

Davide Calandra  
davide.calandra@polito.it

Alberto Cannavò  
alberto.cannavo@polito.it

<sup>1</sup> Politecnico di Torino, Dipartimento di Automatica e Informatica, Corso Duca degli Abruzzi 24, 10129, Turin, Italy

yet). In fact, when targeting subjects with limited skills, AR proved so far to be mostly suitable for use cases that are based on well-established concepts and that do not change for a reasonable time [36].

To deal with this limitation, AR has been largely investigated in combination with remote assistance [1, 24, 30]. The process innovation deriving from the introduction of the latter methodology can reduce the need to have experts on-site and, hence, save time and costs [44]. A number of applications for providing remote assistance using AR in generic scenarios were proposed in the scientific literature [14, 19, 23]. Many commercial platforms specifically designed to support AR remote assistance in industrial contexts are also available [22, 38, 48, 52].

Even though the use of these tools can be beneficial, the way in which AR-enhanced remote assistance is generally deployed today is characterized by an intrinsic weakness. In fact, the common approach is that the expert guides the operator step-by-step until the end of the procedure, in a kind of “explanation-execution” cycle. This approach translates 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 the skilled resource.

In this paper, an alternative to this fully assisted approach is presented, in which the assistance workflow is reorganized with the aim to reduce the time the expert needs to be remain in the call. The proposed approach splits the assistance in two distinct phases. In the first phase, the expert delivers all the information required to deal with the given issue. AR is used not only to support the explanation but also as a way to let the operator access the above instructions when needed and use them at his or her own pace. As the operator starts to perform the required steps, the expert can leave the call. Should additional help be required, the call with the expert could be re-established in order to unlock the situation and resume autonomous operation. This reorganization is expected to have a negligible impact on the overall time spent by the operator to solve the issue.

To implement the proposed workflow, functionalities for placing AR contents in the form of spatially anchored, self-explaining instructions must be provided to the expert. Contents shall be persistent and chronologically navigable even after closing the call, so that the operator can follow the instructions autonomously, similar to what happens with unassisted AR applications. To the best of the authors’ knowledge, a single research prototype or commercial product offering all the functionalities required to support the devised approach has yet to be proposed. For this reason, along with the approach, a remote assistance platform integrating these functionalities is also presented.

The platform was designed in collaboration with KUKA Roboter Italia Spa<sup>1</sup>.

The proposed approach was compared in both objective and subjective terms with commonly adopted, step-by-step AR guidance through a user study that encompassed three different industrial use cases. For a fair comparison, the developed platform was endowed also with functionalities required to implement the traditional approach. Experimental results confirmed the effectiveness of this new approach in reducing the time involvement for the experts with minimal to no impact on the operators’ performance, showing the conditions in which the greatest advantages could be envisaged.

## 2 Related works

In this section, various scientific works and commercial platforms for AR-enhanced remote assistance are analyzed, with the aim to present the common functionalities available in existing applications and identify a set of features, either available or missing, that could help to increase the operators’ autonomy.

Early examples of AR tools proposed in the literature for remote assistance applications are the laser pointer [41], a visual tool that allows the expert to point a specific target while assisting the operator, 3D shapes (arrows, boxes, and circles [40]), texts [2], and hand drawings [40]. AR contents were either inserted on frames captured from the video feed and sent back to the operator [40], directly provided on the video feed without the need to exchange images [5], or attached onto real objects using 6-DOF positional tracking, e.g., by leveraging marker detection [5]. Although with markers it is possible to obtain robust and precise tracking, using this technology alone requires to always frame at least one marker to correctly estimate the camera position and anchor contents [37]. Alternative marker-less technologies for anchoring AR contents were therefore investigated [11, 14]. 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 simultaneous localization and mapping (SLAM), lowering down the costs.

Although all the systems cited 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 (e.g., in the form of a timeline). Moreover, they do not provide the recording of the visual and textual information created

<sup>1</sup>KUKA Roboter Italia Spa: <https://www.kuka.com>

during the session. These data could be possibly reused by the expert and/or made available for future consultation by the operator. Hence, the expert may be required to reiterate the same instructions in case of repetitive steps or should a different operator need assistance on the same or similar topic. Similarly, the operator may need to contact the expert again for problems already solved in the past. The idea of introducing an ordered list of received instructions was explored in [39]. The proposed system 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 [35], an AR framework capable to record data about the intervention while the expert is providing instructions to the operator was presented. The serial number of the object of the assistance is recognized by scanning a QR code, so that relevant information (specifications, history of interventions, etc.) can be retrieved from an archiving server. Video stream is also captured by the device, and real-time object detection is used to identify the printer parts. The data that remain available to the operator after the end of the session can be either 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.

Another example of how contents can be reused is given in [21]. The work presents a framework for the creation of AR-based applications aimed to improve collaboration and support industrial technicians in two different use cases. The first use case focuses on providing technicians with real-time data related to the machines on which they are performing maintenance or repair operations. Augmented contents are shown on HoloLens glasses and can be stored in an external server to enable persistence between sessions. The second use case analyses the provision of remote support by leveraging AR hints on mobile devices. AR features include “billboard” elements (i.e., virtual contents continuously re-orienting towards the user) and the possibility to draw and share 2D sketches directly on the video stream. As suggested by the authors among future developments, the two use cases could be possibly combined in a single solution; in this way, the operator could work autonomously by reusing AR information and anchored contents throughout different sessions, request remote assistance only when needed, and keep working with the received AR instructions after dismissing the expert.

As seen, applications for AR-supported remote assistance proposed in the literature are characterized by a high level of heterogeneity regarding devices, tracking techniques and functionalities made available to the expert, among others. Commercial applications, on the other hand, are lately converging to a common and general-purpose configuration, while trying to encompass all the relevant

use cases for remote assistance. In fact, the wide-spreading of consumer AR devices (like smartphones and tablets) favored the adoption of a common set of functionalities. Thus, e.g., 6-DOF (degrees of freedom) sensor-based positional tracking [29, 43, 45, 49] widely replaced the classical 2D overlaying of AR content (0-DOF, with no tracking) as well as the 6-DOF marker-based approaches [33]. Furthermore, the increase of speed and reliability of mobile networks, the great availability of cloud solutions, and the transition of many business entities to the Software-as-a-Service paradigm led to the appearance of many off-the-shelf remote assistance platforms offered as subscription services.

Besides AR tools, common features offered by these platforms include user registration and authentication, assistance management and scheduling, as well as session recording and archiving [38, 47]. Platforms also offer less usual functionalities, e.g., to support preliminary troubleshooting phases that can be used for known issues not requiring the assistance of an expert. These features are made available through dedicated interfaces for remote experts (web portals or desktop applications [22, 38]) and through AR applications for operators (e.g. for smartphones, tablets, or HMDs).

Some of the commercial products are not implemented as platforms, but rather as stand-alone AR applications that can be symmetrically used by the expert and the operator. In this case, both of them are provided with the same functionalities, since a dedicated portal is not present. Examples of this kind of applications are provided, e.g., in [28] and [46].

Regarding methods used by commercial products to convey information from the expert to the operator, it can be observed that, in most of the cases, instructions are provided using voice (audio-video call), image sharing [28, 38], or instant messaging [38]. That is, AR is mainly exploited to enhance the communication potential of the audio-video call. This observation is confirmed by the fact that the most common AR tools in these products are temporary, i.e., they remain visible only for a limited time, or not self-explanatory, such as hand drawings [45, 49], pointers [38, 43], or anchored shapes (arrows, circles, etc.) [20, 46]. In some cases, anchored texts [47, 48] and images [16, 29] are supported too, but the operator is not provided with a timeline of the delivered instructions, making it very hard for him or her to recognize the order of the received hints and, thus, to easily access them when/as needed. Moreover, the usual feature set of the available products does not include a way for the operator to retain access to the AR experience after closing the call with the expert, nor to preserve the anchored contents for future sessions. In some cases, the operator may retain some sort of session history after the call (the audio-video stream recording, the chat log,

or the shared images, but AR contents and their anchored positions are usually not saved in the recording.

The set of analyzed commercial products and their relevant features are reported in Table 1.

Based on the analysis of scientific works and of commercial products, it was realized that many of the AR features that could support a shift from a step-by-step guidance to a more autonomy-oriented approach have been developed already, but they are not being used yet to this purpose. One of the reasons could be that a single platform integrating them is not available, neither in the literature nor on the market. Thus, in order to study and compare the effectiveness of the two approaches, such a platform had to be developed too.

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/clarity 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.

### 3 AR-supported remote assistance platform

As said, the goal of the present paper was not to develop “just another platform” for remote assistance. Rather, it was to study how the unexpressed potential of AR technology could be exploited to build a resource-effective approach to the problem. The aim was, in particular, to maximize the quality of the support offered to on-site operators while optimizing also the time investment for remote experts.

Hence, a set of requirements and must-have features supporting the above approach were first identified; then, since none of the existing solutions analyzed in the review of state of the art integrated them all, a platform was developed to remedy this lack. Its architecture is reported in Fig. 1. Three main components can be identified: the operator-side mobile application and the expert-side web portal, representing the platform front-end, and a back-end supporting key services such as information exchange, archiving, etc. Design choices and implementation details will be discussed in the following.

#### 3.1 Operator-side mobile application

A number of design choices had to be made concerning, among others, the AR-enabled devices to be supported by the mobile application, the tracking method to adopt, the AR framework to use, and the AR features to exploit.

**Table 1** Features of the main products for AR-based remote assistance

Feature	Product														This work
	[45]	[43]	[49]	[28]	[52]	[46]	[38]	[53]	[6]	[16]	[20]	[47, 48]	[22]	[29]	
Audio-video call	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hand drawings	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
(Laser) pointer	✓	✓			✓		✓								✓
Image sharing				✓	✓		✓		✓	✓	✓	✓	✓	✓	✓
6-DOF positional tracking	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
Anchored shapes (arrows, circles)	✓	✓				✓	✓			✓	✓	✓	✓	✓	✓
Anchored texts	✓	✓				✓				✓	✓	✓	✓		✓
Anchored images										✓				✓	✓
Expert portal (assistance management and scheduling)	✓	✓	✓				✓	✓	✓		✓	✓	✓		✓
Call recording (audio and video)	✓			✓	✓					✓	✓	✓			✓
AR instructions timeline (for the operator)															✓
Past content reuse (for the expert)		✓					✓				✓	✓	✓		✓
Preliminary troubleshooting												✓	✓		✓
AR contents preserved after closing the call															✓
AR contents preserved between calls														✓	

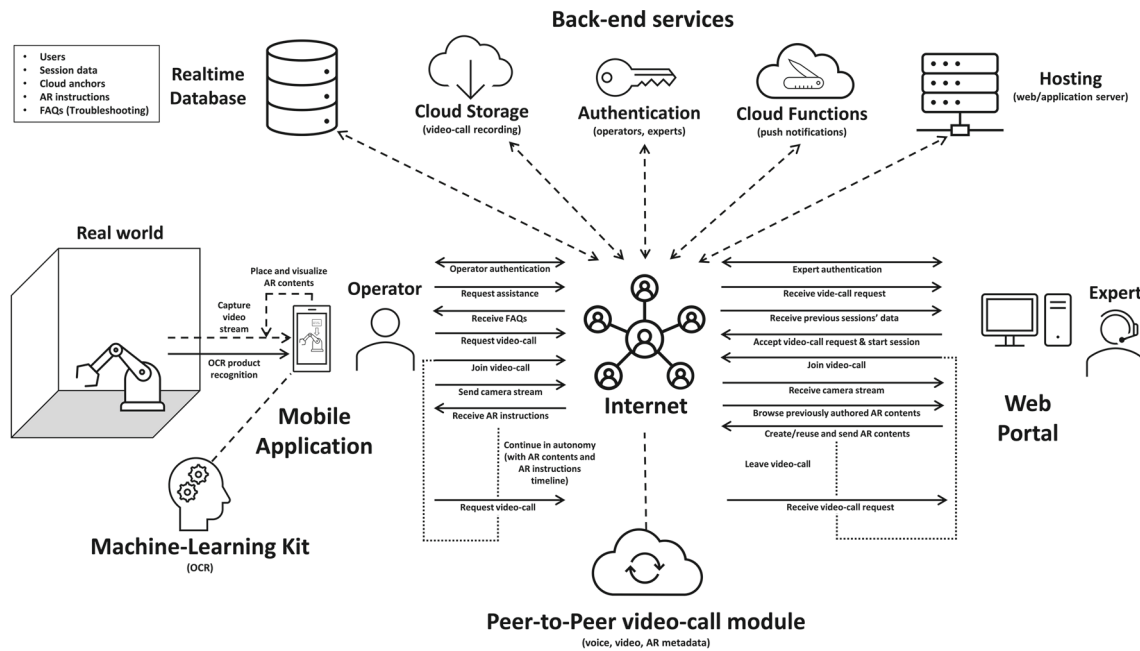


Fig. 1 Architecture of the remote support platform

### 3.1.1 AR devices

For what it concerns the devices to be supported, different alternatives were considered. Although wearable technologies (such as HMDs or smartglasses) may guarantee a higher degree of freedom compared to hand-held devices, previous works proved that they could introduce limitations in terms of interaction; hence, the latter solution is often preferred by the users [54]. Moreover, smartphones and tablets are much more common and wide-spread than wearable devices, and they were shown to be characterized by a higher ease of use as users are already accustomed with them [4, 5, 21, 35]. Finally, the possibility to avoid the use of additional, expensive hardware may be a significant advantage from the viewpoint of the company offering the assistance service, which can serve a wider and more heterogeneous set of customers.

Based on these considerations, the operator-side application was targeted to mobile devices, even though the proposed approach could be also implemented onto HMDs or smartglasses. Given the larger number of devices supporting Android, it was implemented for this environment, but similar functionalities could be provided also in other environments.

### 3.1.2 AR Tracking

Another key point in the implementation of the operator-side application was the selection of the tracking method. In fact, marker-based approaches were proven to be the

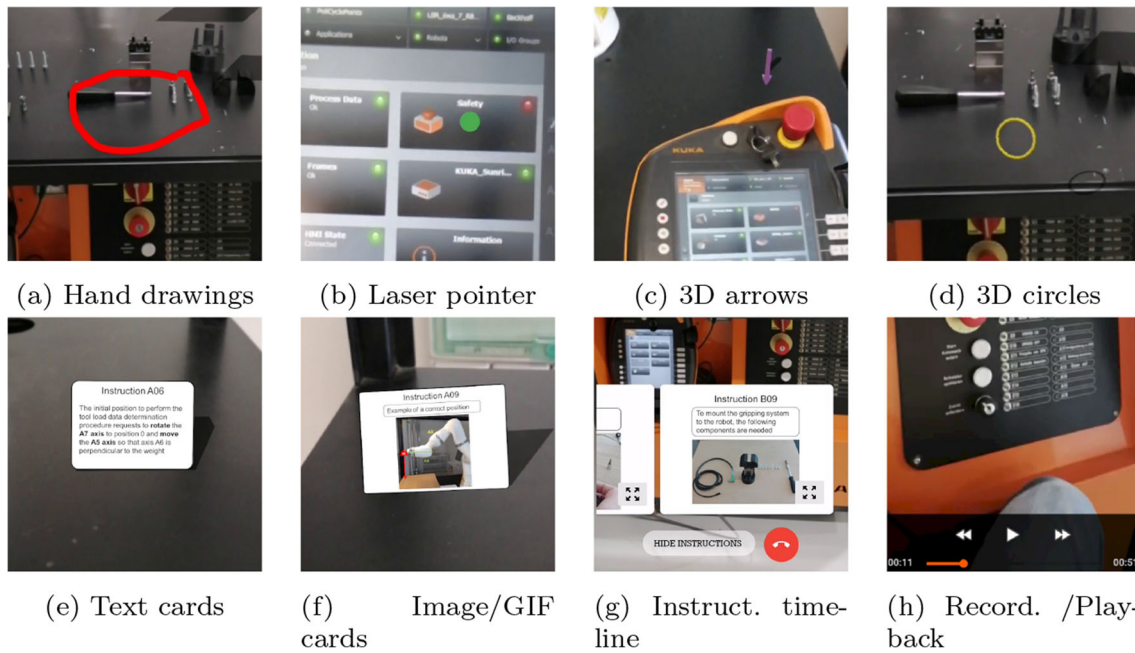
most effective solution for many industrial applications [7, 17]. However, these approaches are characterized by some limitations which are particularly critical for remote assistance, as reported in [15, 34], and [37]. Unlike other “planned” activities such as maintenance or training, remote assistance may not rely on a previous setup phase to place, e.g., markers where necessary. Moreover, the flow of the assistance could be very unpredictable: the remote expert may be requested to devise new solutions to a given problem directly while he or she is providing the support, resorting to methods for quick contents (authoring and) placement.

Indeed, marker or, more in general, image detection can be useful in the considered context to identify known products (and their parts) when this information has been already integrated in the product itself by the manufacturer. For instance, in the developed application, OCR (optical character recognition) was used to identify, before starting the call, the product to be serviced via its product label (similarly to what was done in [35] using a QR code).

However, for the remote assistance session, the marker-based approach was discarded in favor of 6-DOF optical-inertial positional tracking [33]. As seen, most of the commercial solutions use this technique to manage AR contents, since it allows to attach them directly to real-world elements.

### 3.1.3 AR framework

As said, the mobile application was implemented for the Android environment. To develop the AR functionalities,



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

the ARCore<sup>2</sup> library was selected because of its deep integration with the underlying system and its cost-effectiveness. Its 6-DOF positional tracking was exploited to attach AR contents to elements in the operator's space, and its “anchoring” feature was used to keep the above contents in place during the assistance session. Within a session, a call with the expert can be established, closed, and then re-established as needed; through anchoring, AR contents are preserved also throughout these calls. The devised assistance paradigm can thus provide operators with helpful information that could increase their ability to operate autonomously. Hopefully, this will reduce the load on the remote experts for solving the problem in that specific situation, but also should that situation occur again in the future.

### 3.1.4 Supported features

The client application supports registration/authentication, product recognition (via OCR on the product label, if available) and initial troubleshooting. If the operator is not able to solve the problem autonomously with the provided frequently asked questions (FAQs), then the application lets him or her initiate an AR-supported remote assistance session (or schedule it).

An audio-video call is set up. During the call, the mobile device's camera is used to provide the expert with a video feed about the serviced product and the surrounding

environment, whereas the screen is used as a “magic-window” for displaying AR contents to the operator.

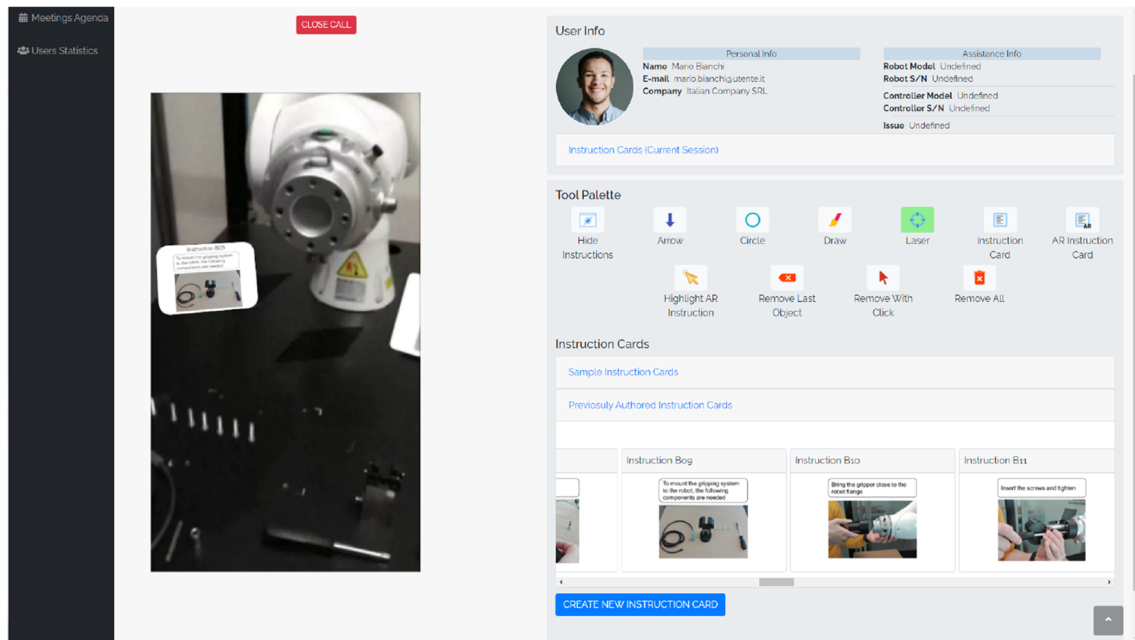
The expert can provide support to the operator using all the most common AR-based tools identified in the review of the state of the art, whose implementations are shown in Fig. 2. The tools can be split in two categories, i.e., temporary and persistent.

Temporary tools include hand drawings (Fig. 2a) and laser pointer (Fig. 2b), and rely on graphics contents which are displayed as 2D overlays over the video feed; according to [33], these 0-DOF elements can be referred to as “0D” AR contents. These elements are non-permanent, i.e., they vanish few seconds after placement. Persistent tools, on the other hand, exploit 6-DOF tracking to let the expert place 3D shapes like arrows and circles (Fig. 2c and Fig. 2d) as well as instruction cards (Fig. 2e and Fig. 2f); these contents are attached to real-world elements using spatial anchoring (hence, in [33], they are named “6D” AR contents). Cards, in particular, can either contain text, images, or animated GIFs. If cards do not require to be spatially anchored in the real world, they can be just inserted as 0D contents.

A key feature of the devised platform is the possibility to record the AR contents exchanged during the session. Once placed, the contents are displayed in a timeline within a scrollable panel located in the lower part of the interface (Fig. 2g) which can be shown/hidden by pressing a button.

Not all the contents are stored in the timeline. Very temporary or explanation-complementary tools (laser pointer, hand drawings, and 3D shapes) are not considered, although some of them will be spatially persistent throughout the

<sup>2</sup>Google ARCore: <https://developers.google.com/ar>



**Fig. 3** Interface of the expert-side web portal

session. On the contrary, text/image cards, which represent powerful tools for the expert to visually fix in the operator's mind the concept he or she is explaining using the voice, are recorded for later use, being them anchored or not. These contents can be accessed both within the current session as well as in the future. The timeline panel can also be used to enlarge (full-screen) or reduce 0D cards, as well as to highlight the actual card to be displayed in case of multiple occluding 6D cards.

The augmented video-stream is also saved (with bookmarked instructions) and made accessible to the user for a full recap of the session (Fig. 2h).

### 3.2 Expert-side web portal

The expert-side of the platform was developed as a web portal, similar to [14]. The portal's interface is shown in Fig. 3.

During the call, the interface provides the expert with information concerning the operator (company, equipment, etc.) and previous requests for assistance (if any). Data collected by the operator through the initial troubleshooting phase are also reported, offering the expert information that could help him or her to frame the context of the assistance request and possibly anticipate operator's needs.

Similar to [45], the Twilio<sup>3</sup> APIs were used to establish a peer-to-peer bi-directional audio and mono-directional (operator-to-expert) video communication with

the operator's side. Video feed is displayed in the portal's interface. Since Twilio APIs support the exchange of other data between the involved peers and can be integrated with ARCore to create collaborative AR experiences, they were also exploited to support the transfer of AR data.

AR tools that can be used by the expert to support voice explanations are grouped in a palette displayed in the portal's interface. The expert can control the laser pointer or make hand drawings appear on the screen of the operator's device by using the mouse on the received video feed.

He or she can also choose the 3D shapes to be added in the operator's space. While the expert places them on the video feed, the operator-side application tries to attach them to real-world objects by estimating planar surfaces in the camera's field of view using ARCore. Hopefully, added shapes will be displayed in the same place independent of operator device's movements.

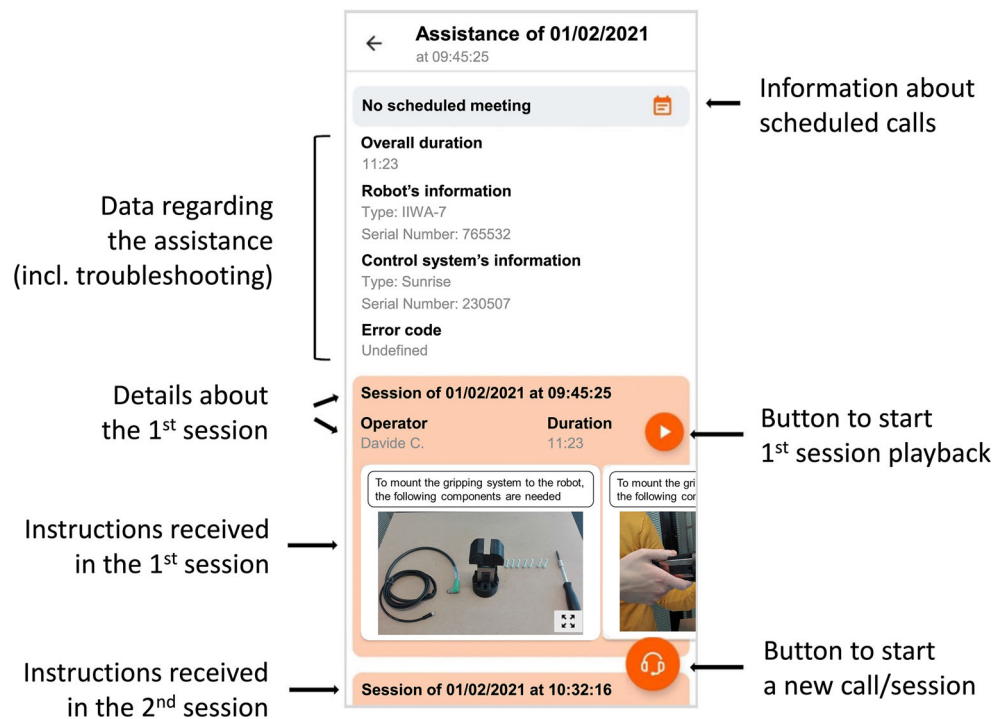
The expert can also add text/image instructions that, as said, are displayed in a chronological order as scrollable cards in the lower part of the operator-side application's interface (as well as in the web portal's interface); the instruction cards can be either displayed as 0D elements or anchored to the real world as 6D elements. 6D cards' tilting is controlled so that they are always oriented for best readability.

Cards can be either picked up from a list of common instructions, selected from those used in a previous session, or created on-the-fly for the specific session. As said, cards can contain text or (animated) images. Similar to what happens on the operator's side, the expert can select added

<sup>3</sup>Twilio APIs: <https://www.twilio.com/>



**Fig. 4** Session handling on the operator-side application



instructions (by clicking them either in the list or in the video feed) to highlight them, e.g., to catch the operator's attention during the explanation or resolve occlusions.

As said, the expert and the operator can close the call at any time, but they will retain the possibility to browse and visualize previously created AR instructions. Should the operator need further help, he or she could request a new call; in this case, previously placed contents are supposed to retain the original position. Finally, the operator is provided with a session history, through which the recap of previous sessions (instructions timeline and audio-video recording) can be visualized for future reference.

### 3.3 Back-end

A key characteristic of the remote assistance paradigm supported by the devised platform is the possibility to have sessions that can be connected each other and restored when necessary. In this way, the remote expert can, e.g., leverage instructions delivered in previously completed sessions for the same or similar problems, whereas the operator can retrieve instructions from a previous session to execute a procedure for which assistance was received in the past without asking for further support.

To this aim, platform's back-end development was centered on the concept of session, and both the mobile application and the web portal rely on this concept for operation.

Back-end was implemented using Google Firebase<sup>4</sup>, leveraging some of its off-the-shelf features such as Authentication, Realtime Database, Cloud Storage, ML-kit, Cloud Functions, and Hosting. Building on them, a networked platform was created, supporting user registration, authentication, call scheduling, session management/archiving, recording, push notifications, and OCR (for the troubleshooting).

For each session, the platform records information collected in the troubleshooting phase as well as instructions provided by the expert using available AR tools (Fig. 4). The same information is displayed also in the web portal's interface. Independent of the time passed between two sessions and of who actually provided the support, the expert has at his or her disposal helpful information that can ease the identification/solution of the problem.

Session recording is performed on the web portal's side, but storage is handled in the back-end. Thus, recorded sessions can be made available also on different devices. The communications of both peers are saved, together with the video feed received by the mobile device side and graphic contents drawn on it. Bookmarks are also set, allowing to quickly jump to frames where instructions were provided using AR tools.

<sup>4</sup>Google Firebase: <https://firebase.google.com/>

## 4 Experimental setup

As said, the peculiarity of the remote assistance approach proposed in this paper lays in the possibility for the remote expert to provide persistent, self-explanatory AR instructions that can be used by the operator to work autonomously after the end of the call. To assess the effectiveness of this approach, a platform was first developed to support it. Afterwards, a preliminary study was carried out with 23 volunteers with the aim to validate the usability of the developed platform and design the evaluation methodology. Finally, the platform was used to compare the proposed approach with step-by-step guidance through a user study in which 60 volunteers were involved in three different use cases encompassing a collaborative robot.

### 4.1 Participants

The 60 volunteers (49 male and 11 female) were aged between 19 and 66 ( $\mu = 32.86$ ,  $\sigma = 9.54$ ). Some of them were from the administrative staff of KUKA Roboter Italia Spa; remaining subjects were recruited among students and academic staff at the authors' university. Informed consent was obtained. According to information collected with a demographic questionnaire, participants reported a good knowledge of the Android environment and a medium knowledge of audio-video conferencing applications. Moreover, they stated to be poorly familiar with the concept of AR, and reported a very low previous experience with AR applications and with tools (both software and hardware) for controlling collaborative robots.

### 4.2 Methodology

Participants were first requested to fill in a demographic questionnaire aimed to evaluate their knowledge of technologies used in the experiments, and their previous experience with them. Then, they were introduced to the experimental material, focusing on functionalities needed to perform the operations requested by the three use cases. In particular, details about the mobile application, the collaborative robot, the SmartPad (the device used to manually control the robot), and the other tools required during the experience were presented. Participants were given time to familiarize with the SmartPad, with the smartphone used to run the mobile application (a Huawei HONOR 8X), and with the application itself. Finally, they were introduced to the experiment's goals.

During the experiment, each participant played the role of a generic operator needing support to perform a specific procedure on the robot (detailed in Section 4.3).

Hence, they were requested to launch the application for remote assistance on the mobile device, complete the troubleshooting phase by collecting information regarding the robot, and initiate the video call for receiving support. A KUKA technician in charge of customer service played the role of the remote expert responding to the call and managing the web portal.

The assistance could be provided in two different modalities. In the first modality, named *fully assisted*, or *FA*, the expert provides continuous support to the operator through voice/video conferencing and AR-based tools until the problems is solved; instructions are provided step-by-step as the operator proceeds through the procedure and new issues actually come up.

In the second modality, the expert assists the operator until a certain point of the procedure, then asks him or her to proceed autonomously. From this point on, the operator needs to make a larger use of received instructions, particularly of anchored text/image cards that are controlled (activated/deactivated and highlighted) through the chronologically ordered, scrollable list. If needed, he or she may also call the remote expert again: in this case, the expert may leverage both new and previously provided instructions in order to help the operator to solve faced issues. Since this modality was designed to improve operator's ability to act autonomously (at least for part of the task), it will be later named *partially assisted*, or *PA*.

For each considered use case, participants were split into two groups: half of them were asked to carry out the task in the FA modality, the other half in the PA modality. Assignment to groups was made by trying to balance as much as possible distribution in terms of age, gender, and previous experience.

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, hand drawings, and 3D shapes; 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, rarely the temporary tools (the laser pointer and 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.

During the experiments, objective data on participants' performance were collected. At the end of the experiment, participants were asked to fill in another questionnaire aimed to evaluate their experience in subjective terms.

**Fig. 5** KUKA's LBR iiwa 7 R800 collaborative robot



Objective and subjective metrics will be discussed in Section 4.4.

### 4.3 Use cases

In order to compare participants' performance and evaluate their experience, three tasks were selected among the most common and relevant remote assistance procedures regarding collaborative robots. The aim was to isolate possible impacts of the given task on the effectiveness of the assistance approach being used. All the tasks involved a KUKA LBR iiwa 7 robot (Fig. 5). The robot was equipped with an interactive flange (Media flange Touch pneumatic) that allows manual jogging by means of an enabling switch on the flange itself. Several videos showing an example of assistance with the two modalities for each use case are available for download<sup>5</sup>.

<sup>5</sup>Videos of the two assistance modalities for the three use cases: <http://tiny.cc/pf87qz>

#### 4.3.1 Gripper assembly (GA)

In this use case, the remote expert provides assistance regarding the assembly of the various components of a robotic gripper (based on the Schunk EGP 40-N-N-B<sup>6</sup> gripping system, shown in Fig. 6) on the robot flange. Elements to be handled include the gripping system, a pair of custom 3D-printed gripping fingers, a custom 3D-printed connection flange, and three sets of screws along with a screwdriver with interchangeable heads (one for each type of screw). The steps to assemble the gripper can be summarized as follow.

1. A gripping finger has to be inserted on the base jaw of the gripping system, then fastened with a screw; this step had to be repeated for the finger on the other jaw.
2. The gripping system has to be inserted into the connection flange, which is used as interface between

<sup>6</sup>Schunk EGP 40-N-N-B: <https://schunk.com/us.en/gripping-systems/product/28088-0310940-egp-40-n-n-b/>

**Fig. 6** Schunk's EGP 40-N-N-B gripping system with custom fingers and connection flange disassembled (left), and ready to be mounted (right)



- the gripper and the robot flange, and fastened with two other screws.
3. The resulting assembly has to be mounted on the flange using seven screws.
  4. A cable has to be used to connect the gripping system connector (4-pin M8) with the X3 connector on the robot flange (17-pin M8).

In a traditional remote assistance call, this task would imply a lot of downtime for the expert, mainly due to the tightening of the many screws. The actual downtime may differ from one operator to the other based on their manual skills and previous experience with this fairly common kind of task.

#### 4.3.2 Load data determination (LD)

In this procedure, the robot executes multiple measurements with different orientations of the wrist axes. These runs determine the mass and the position of the center of mass of the tool mounted on the robot flange. The robot has to be first moved to a specific position. In particular, the seventh axis has to reach the zero position, whereas the fifth axis has to be rotated so that the sixth axis is perpendicular to the weight. The key steps of the procedure are reported below.

1. The user moves the robot to a valid position, considering constraints above.
2. The user navigates the SmartPad interfaces to the Robot view, and selects the Load data button.
3. In the Load data view, the user selects from the Tool selection list the tool for which the load data has to be determined.
4. From this point on, the user has to press and hold down the enabling switch until measurements have been completed; while holding down the enabling switch, the user has to press the Determine load data button.
5. If a previous measurement already exists for the selected tool, the user can choose the option Redetermine mass to recalculate the data.

6. The load data determination consists in a predefined set of robot movements involving the fifth and seventh axis; during this process, a progress bar is displayed on the SmartPad.
7. At the end of the process, the determined load data are displayed, and the user has to press the Apply button to save and use them.

This task, which is a routine configuration activity for this kind of robots, is characterized by many simple interactions with the SmartPad, e.g., for navigating the menus (via touch interactions) or jogging the robot axes (via physical buttons) to reach the particular pose required for starting the measurements (Fig. 7). Like for the GA task, this task implies a downtime for the expert, but in this case it is partly related to the operator's ability (in posing the robot), and partly fixed (due to the automatic movements in the measurement process).

#### 4.3.3 Emergency recovery (ER)

This procedure is meant to recover from an emergency stop. It has to be performed when the robot violates one of the safety rules set up in the cell configuration. In the prepared use case, the robot exited the configured safety volume while running a given program. In this condition, the program is automatically blocked and robot jogging is not allowed. To restore the working condition, the following steps need to be executed.

1. The user identifies the error that caused the emergency stop by reading the log in the Safety status window of the SmartPad interface. Additionally, he or she can visualize the violated safety volume boundaries through a web dashboard (KUKA Safety Visualization) accessible through an external PC connected to the robot network.
2. The user turns the key-switch on the SmartPad to show the Connection manager and selects the operating mode

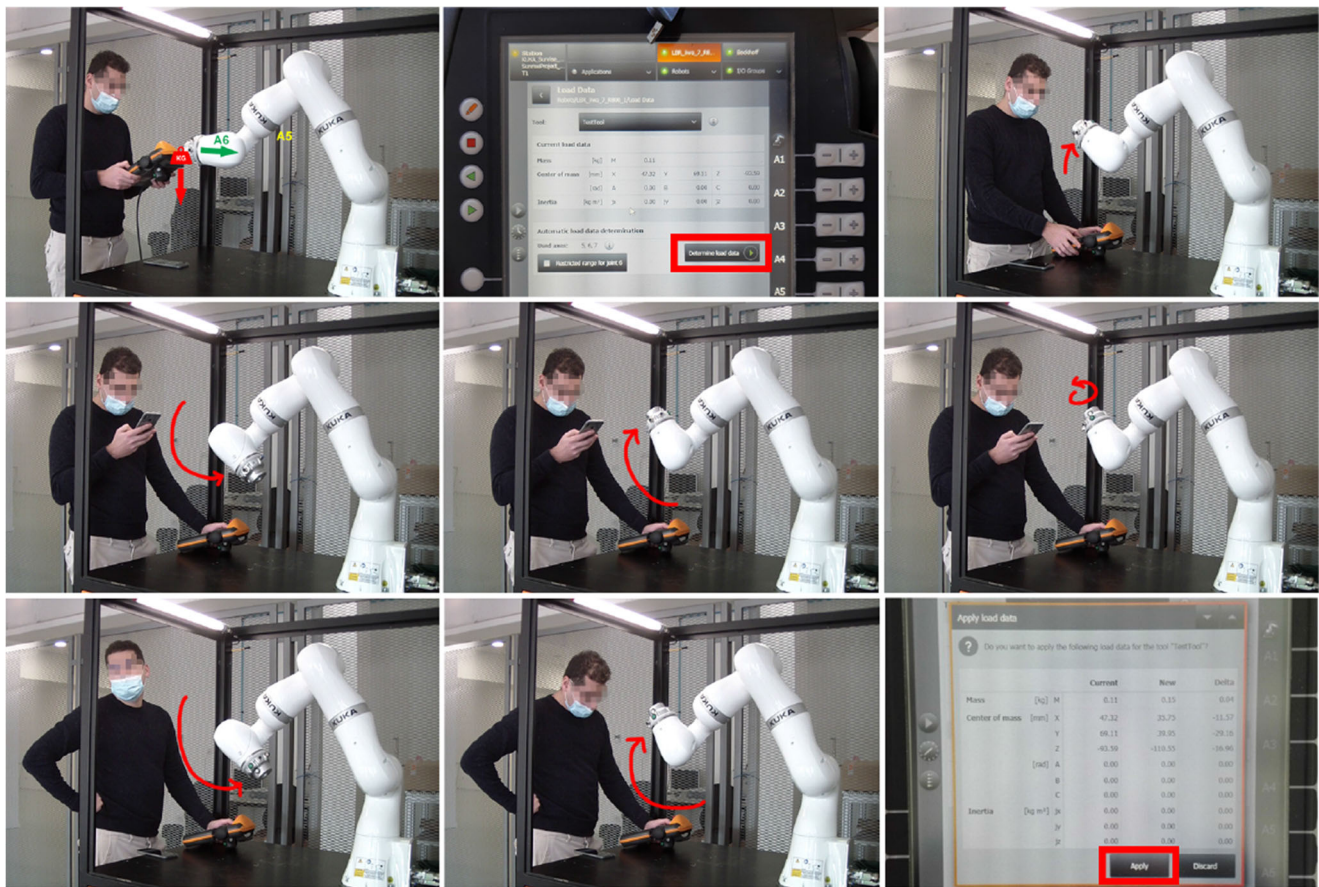


Fig. 7 Launch and execution of the automatic measurements in the LD task

labeled KRF (Kontrollierte roboterfahrt – Controlled robot travelling mode).

3. The user moves the robot in order to make it assume a valid position within the safe volume.
4. Finally, the user verifies if the issue occurs again by re-executing a full cycle of the previously halted program.

The main steps are illustrated in Fig. 8. The robot can be manually moved by handling the flange while pressing the enabling switch on it. The sample program considered for the experiments was a cyclic trajectory composed by eight PTP (point to point) motions between three predetermined points.

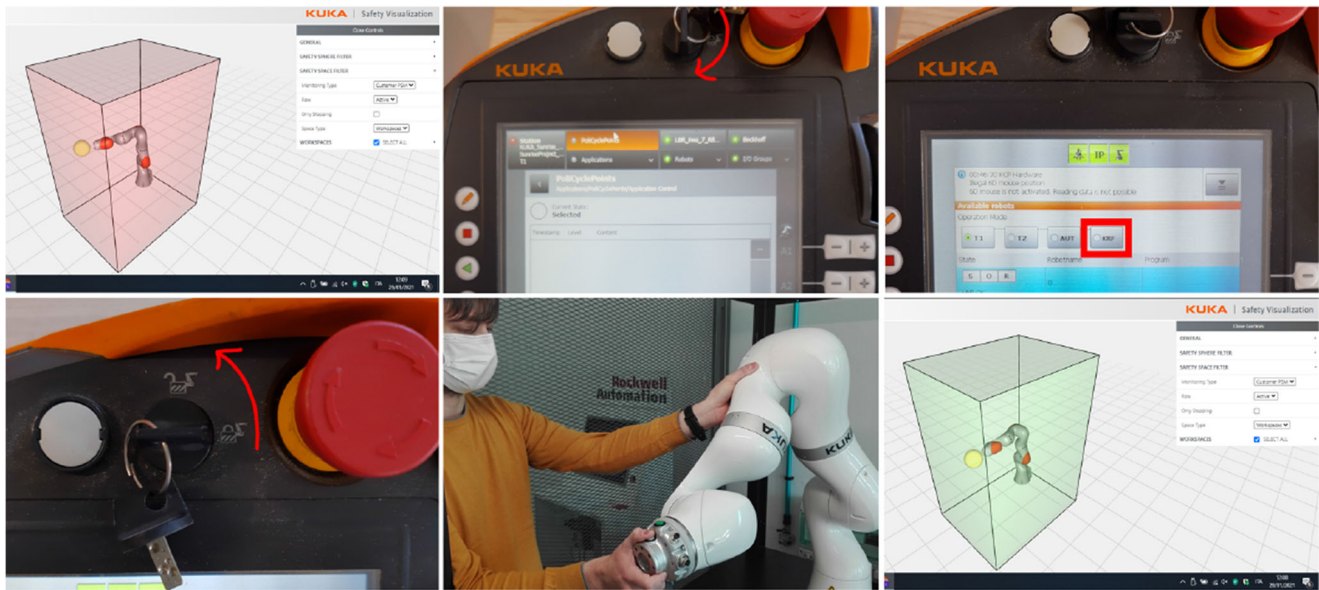
This task is characterized by simple operations, like touchscreen interactions and direct robot manipulations but, differently than the LD task, requires the understanding of some specific theoretical concepts related to the emergency status (the safe volume and the KRF mode) which may be less familiar to inexperienced users.

#### 4.4 Evaluation criteria

As anticipated, participants' performance and experience were evaluated in both objective and subjective terms.

For what it concerns objective evaluation, two metrics were collected. The first metric, named *call time*, corresponds to the overall duration of the communication between the operator and the expert. The second metric, named *completion time*, accounts for the time needed to complete all the steps of the procedure. In the FA modality, call time corresponds to completion time, since operator's actions are supervised until the end of the task; in the PA modality, it corresponds to the cumulative duration of the assistance calls made by the operator during the experiment. For the PA modality, the *number of re-calls* was also recorded.

Subjective evaluation was performed by asking participants to fill in a post-test questionnaire (available for



**Fig. 8** Operations for unblocking the robot in the ER task

download<sup>7</sup>). The questionnaire included specific statements to assess perceived user performance, system performance (in term of quality of the audio-video communication), learnability, memorability, and frustration. The two closing statements were used to assess the suitability of a given assistance approach to the specific task. Each statement was evaluated on a 5-point Likert scale. Space for comments was also provided.

## 5 Results

Results obtained by applying the metrics presented in the previous section were used to compare the two modalities.

### 5.1 Objective results

Measurements concerning completion time and call duration for the three tasks are reported in Fig. 9. Unpaired samples *t*-tests with 5% significance ( $p < 0.05$ ) were used for the statistical analysis.

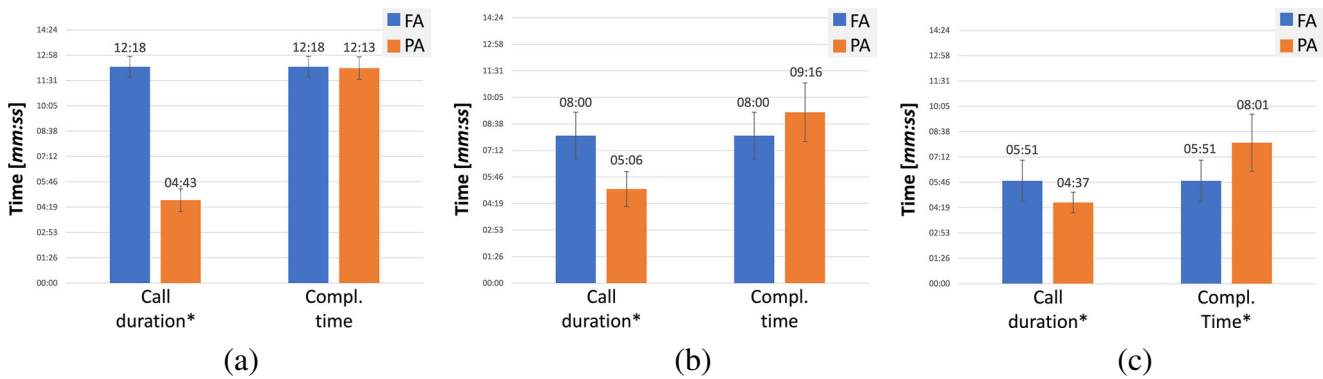
Regarding call duration, it can be observed that with the PA modality, the time invested by the expert was 61.65% ( $p < 0.0001$ ), 36.25% ( $p < 0.0001$ ), and 21.08% ( $p = 0.0109$ ) shorter than with the FA modality in the GA, LD, and ER tasks, respectively. GA is the task in which the advantages of the PA modality were more evident. This result was probably due to the presence of steps in which the expert had to wait for the operator

to complete long-lasting operations (e.g., tightening the screws). The LD and ER tasks were also characterized by steps requesting the expert to wait for the operator, but these steps were related to far less familiar concepts; hence, they required more explanations from the expert in order to be executed. This aspect reduced the advantages associated with the PA modality, since it increased the percentage of the session time used by the expert to illustrate the procedural operations. This fact is particularly critical when the time actually required to perform the considered steps is shorter than the time invested to provide instructions for autonomous operation.

Concerning completion time, no statistically significant difference was observed for the GA and LD tasks between the FA and PA modalities. These results suggest that, in general, the paradigm shift introduced by the devised approach did not impact on the overall operators' performance. However, in the ER task, subjects were 27% faster with the FA modality than with the PA modality ( $p = 0.0044$ ). This finding suggests that, in this kind of task, the proposed paradigm can have a negative impact on the operators' performance. Although the LD and ER tasks apparently encompass similar operations (jogging the robot and navigating the SmartPad interface), it was observed that the latter required the subjects to rely on a larger number of concepts to properly execute the steps in complete autonomy (e.g., to understand if operations had been correctly performed or not).

The different performance of the two modalities in the three tasks can be also observed by considering the number of re-calls in Table 2.

<sup>7</sup>Questionnaire: <http://tiny.cc/4ewqtz>



**Fig. 9** Objective results concerning call duration and average completion time (standard deviation expressed through error bars) collected for a GA, b LD, and c ER tasks. Significant results are marked with the \* symbol

## 5.2 Subjective results

For what it concerns the subjective evaluation, results are reported in Table 3. Statistical significance was tested with the same methodology adopted for the objective measures. In the following, the analysis will focus first on the overall results, tabulated in column “All tasks” (calculated by averaging values on the three tasks); afterwards, the discussion will consider results for individual tasks (remaining columns).

Starting from the statistically significant differences concerning learnability, subjects found the PA modality to be more effective than the FA modality in conveying helpful information about both theoretical concepts (statement 5, 3.87 vs 4.50,  $p = 0.0200$ ) and practical operations (statement 6, 4.13 vs 4.73,  $p = 0.0127$ ), thus making these contents easier to learn. With the FA modality, subjects were guided by the remote expert step-by-step, and were not actually expected/requested to memorize/understand what they were doing. On the other side, the PA modality was explicitly designed to make the subjects work autonomously: hence, they were more solicited to memorize/understand the operations to be performed. As a further confirmation, subjects who used the FA modality felt the need to learn the information required for completing the task (statement 7) more than PA ones (2.33 vs 3.27,  $p = 0.0064$ ). These results are also in line with the scores assigned to statement 8, which indicate that subjects had the impression to be performing the operations

more “mechanically” (i.e., without understanding the real reasons) with the FA modality than with the PA modality (3.50 vs 2.17,  $p = 0.0004$ ).

Results concerning memorability indicate that this modality also made the subjects more confident that they could repeat the procedure again both in the short term (statement 9, 3.63 vs 4.23,  $p = 0.0446$ ) and in the long term (statement 10, 2.73 vs 3.50,  $p = 0.0174$ ). This is probably due to the fact that subjects who used the PA modality partially experienced the possibility to work without the support of a remote expert. The same trend, but with higher scores, was observed when asking subjects about their confidence on the ability to complete the procedure again with information remaining on the device after the call, i.e., instructions timeline and audio-video call recording (statement 11, 4.67 vs 4.93,  $p = 0.0092$ , and statement 12, 4.46 vs 4.86,  $p = 0.0125$ ).

Finally, it is worth observing that the received assistance put more pressure on subjects with the FA modality than with the PA modality (statement 13, 1.63 vs 1.17,  $p = 0.0237$ ). This pressure may be related to the constant presence, in the FA modality, of the remote expert, who had to wait until the completion of each instruction performed by the operator (statement 14, 1.57 vs 1.10,  $p = 0.0092$ ).

With respect to individual tasks, as already hypothesized based on the objective results, the GA is the task which benefited more from the proposed approach. For this task, all the significant differences observed in the overall results were confirmed. In addition, the PA modality showed to be significantly better than the FA modality for statement 2 too. This outcome is probably due to the fact that subjects reported that they needed less support when using the PA modality than the FA modality. Comments provided at the end of the experience suggest that this result could be related to the way instructions were delivered in the two modalities. In the FA modality, in which they were delivered step-by-step, it happened that subjects asked the expert for information yet to be provided (e.g., because

**Table 2** Calls made by PA participants for each task

Task	Single call	+1 call	+2 calls
GA	10	0	0
LD	9	0	1
ER	6	3	1

**Table 3** Subjective results for the user study. The higher the score, the higher the agreement

Category	Statement	All tasks		GA		LD		ER	
		FA	PA	FA	PA	FA	PA	FA	PA
User perf.	1. I think that without the AR hints my performance would have been worse	4.00	4.37	3.80	3.90	4.00	4.60	4.20	4.60
		<i>p</i> = 0.1352		<i>p</i> = 0.8363		<i>p</i> = 0.1768		<i>p</i> = 0.2457	
	2. I did not need any further help from the expert when performing the given task	4.23	4.30	4.10	<b>4.90</b>	4.60	4.40	4.00	3.60
		<i>p</i> = 0.8171		<i>p</i> = 0.0403		<i>p</i> = 0.6278		<i>p</i> = 0.5258	
	3 I would have liked to receive more information from the expert to execute each procedural step	1.53	1.27	<b>1.90</b>	1.10	1.40	1.50	1.30	1.20
		<i>p</i> = 0.1518		<i>p</i> = 0.0261		<i>p</i> = 0.7947		<i>p</i> = 0.6278	
Sys. perf.	4. I thought that the audio-video communication quality was adequate	4.87	4.93	4.90	4.90	5.00	4.90	4.70	5.00
		<i>p</i> = 0.3980		<i>p</i> = 1.0000		<i>p</i> = 0.3305		<i>p</i> = 0.0652	
Learnability	5. The received assistance let me effectively learn required concepts	3.87	<b>4.50</b>	3.20	<b>4.80</b>	4.10	4.40	4.30	4.30
		<i>p</i> = 0.0200		<i>p</i> = 0.0040		<i>p</i> = 0.4762		<i>p</i> = 1.0000	
	6. The received assistance let me effectively learn procedural steps	4.13	<b>4.73</b>	3.70	<b>4.90</b>	4.30	4.90	4.40	4.40
		<i>p</i> = 0.0127		<i>p</i> = 0.0030		<i>p</i> = 0.1033		<i>p</i> = 1.0000	
	7. I felt the need to learn the steps of the procedure and the concepts explained by the expert to complete the task	2.33	<b>3.27</b>	1.40	<b>3.00</b>	2.50	3.40	3.10	3.40
		<i>p</i> = 0.0064		<i>p</i> = 0.0065		<i>p</i> = 0.0782		<i>p</i> = 0.6278	
	8. I performed the operations in a “mechanical” way, without understanding the concepts/reasons	3.50	<b>2.17</b>	4.00	<b>1.70</b>	3.50	2.40	3.00	2.40
		<i>p</i> = 0.0004		<i>p</i> = 0.0003		<i>p</i> = 0.0681		<i>p</i> = 0.4239	
Memorability	9. I think that I would be able to repeat the procedure alone right now	3.63	<b>4.23</b>	4.30	<b>5.00</b>	3.30	<b>4.40</b>	3.30	3.30
		<i>p</i> = 0.0446		<i>p</i> = 0.0041		<i>p</i> = 0.0260		<i>p</i> = 1.0000	
	10. I think that I would be able to repeat the procedure alone in the future	2.73	<b>3.50</b>	3.20	<b>4.70</b>	2.50	<b>3.50</b>	2.50	2.30
		<i>p</i> = 0.0174		<i>p</i> = 0.0006		<i>p</i> = 0.0431		<i>p</i> = 0.6945	
	11. I think that I would be able to repeat the procedure alone right now, using the information available on the device (session history)	4.67	<b>4.93</b>	4.60	<b>5.00</b>	4.60	<b>5.00</b>	4.80	4.80
		<i>p</i> = 0.0092		<i>p</i> = 0.0248		<i>p</i> = 0.0248		<i>p</i> = 1.0000	
	12. I think that I would be able to repeat the procedure alone in the future, using the information available on the device (session history)	4.46	<b>4.86</b>	4.50	<b>5.00</b>	4.30	<b>4.90</b>	4.60	4.70
		<i>p</i> = 0.0125		<i>p</i> = 0.0382		<i>p</i> = 0.0453		<i>p</i> = 0.7486	
Frustration	13. The kind of received assistance put pressure on me	1.63	<b>1.17</b>	1.80	<b>1.00</b>	1.70	1.40	1.40	1.10
		<i>p</i> = 0.0237		<i>p</i> = 0.0131		<i>p</i> = 0.5375		<i>p</i> = 0.2322	
	14. I felt pressure when the expert was waiting for me	1.57	<b>1.10</b>	1.80	<b>1.10</b>	1.50	1.30	1.40	1.00
		<i>p</i> = 0.0092		<i>p</i> = 0.0131		<i>p</i> = 0.5995		<i>p</i> = 0.0871	
Appropriat.	15. I would have preferred to receive assistance in a different way	1.57	1.70	2.10	<b>1.30</b>	1.40	1.30	<b>1.20</b>	2.20
		<i>p</i> = 0.5540		<i>p</i> = 0.0344		<i>p</i> = 0.5708		<i>p</i> = 0.0177	
	16. The received assistance was ideal for the task that I carried out	4.13	4.10	3.70	<b>4.40</b>	4.30	4.40	<b>4.40</b>	3.50
		<i>p</i> = 0.8801		<i>p</i> = 0.0179		<i>p</i> = 0.8061		<i>p</i> = 0.0287	

Bold font is used to indicate the best of the two values in case of significant differences (*p* < 0.05)

included in a following step). This is particularly critical in the GA task, being it characterized by higher familiarity of concepts than the LD and ER tasks. 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. Instructions were also available in the timeline, and these factors could have

contributed at making subjects feel that no further help from the expert was needed.

The ER task, on the other hand, did not show any of the overall significances, making it impossible to discriminate between the two modalities.

Finally, the LD task apparently represented an intermediate case, as significant differences were found only for a



subset of the previously analyzed statements. The ER and LD tasks were probably perceived as less familiar than the GA task, and this fact could have influenced learnability (statements 5–8). Since the GA and LD tasks had a comparable (low) complexity, the advantages of the PA modality in terms of memorability were confirmed (statements 9–12). It is worth observing that results for statement 13 (frustration) were only significant for the GA task, which is characterized by steps with high downtime and made the subjects feel a high pressure in the FA modality because of the presence of the expert. This finding, i.e., the preference for the PA modality that gets less marked when passing from the GA, to the LD and ER tasks, appears to be in line with objective results.

The last two statements (15 and 16), designed to evaluate the suitability of the assistance approach to the given task, provided a further confirmation of the observed trend: the PA modality was perceived as more appropriate than the FA modality for the GA task and less appropriate for the ER task, whereas nothing can be said for the LD task.

To summarize, in the considered use cases, the PA modality proved to be capable of reducing the time invested by the expert, with either a positive or a negligible impact on operators' performance. In fact, subjects involved in the experiments generally completed the assigned task in a comparable time with the two modalities, and preferred the PA modality for many of the explored subjective dimensions.

## 6 Discussion and conclusions

This paper proposed an approach to reduce the time invested by a remote expert while providing support to on-field operators through AR-powered remote assistance tools in the industrial field. The idea behind the devised approach is to deliver and discuss with the operator all the required instructions in bulk at the beginning of the assistance, then let him or her execute the operations autonomously until the problem is solved.

To evaluate the possible advantages brought by the adoption of this approach, which in the paper has been referred to as partially assisted (PA), a comparison with a fully assisted (FA) approach in which the expert provides continuous, step-by-step support from beginning to end was performed by means of a user study that considered three different industrial tasks.

Results showed that the PA approach significantly reduced the time of the expert intervention in all the considered tasks, allowing the operators to successfully complete the procedure in an autonomous way. The advantages of PA were more evident in the tasks characterized by many steps

and encompassing important downtime for the expert. In most of the tasks, time requested by the operators to complete the operations did not differ significantly in the two modalities, except in one task in which the explanation was particularly complex and the operations were then relatively quick to execute. This is the only task in which the PA approach was not preferred to the FA approach by subjects involved in the experiments. In general, subjects perceived the proposed approach as significantly more useful, capable of making them work more efficiently and relaxedly, and to convey the expert's knowledge better than the FA approach. The subjective evaluation on specific tasks showed that the higher was the unfamiliarity of the involved operations and of the complexity of concepts behind them, the lower was the perceived advantages of the PA approach with respect to the FA approach.

It is worth observing that the effectiveness of the devised approach depends on the ability of the expert to create correct and easy-to-understand AR instructions, which allow different operators (who may have different skills/background) to complete the task autonomously. When dealing with unknown issues, the procedure should be defined on-the-fly by the expert; however, for the next requests, it would be possible to leverage the already acquired knowledge and generated AR contents.

As for future developments, several directions could be explored. Currently, a mechanism letting the operator check if the steps in the received instructions have been correctly executed or not is not available. In case of doubts, the only way to verify that aspect is to re-establish the call with the expert. Techniques could be developed to let the system automatically recognize the outcome of each step by using, for instance, object recognition algorithms (e.g., for assembly tasks) or dedicated procedures to monitor real-time data provided by machinery (e.g., for repair task). These mechanisms could also solve issues related to the visualization of a cumbersome amount of instructions all at once (since steps would be gradually revealed). Most importantly, they would contribute at bringing the proposed approach even further, as the assistance workflow could be transformed from a set of sequential instructions to a more complex organization, e.g., exploiting also conditions and branches, suitable to the many variations of the situation faced during autonomous operation.

Finally, it shall be considered that, while the operator is carrying out the assigned procedure, the expert should be ready to re-enter the call in case the operator faces new problems. Should the initially involved expert be unavailable, a mechanism to transfer the request to another expert could be developed, with the aim to limit as much as possible the time required by him or her to understand the context and seamlessly continue the assistance.

**Acknowledgements** The authors want to acknowledge the help given by KUKA Roboter Italia Spa and SCHUNK Intec Srl which made available the equipment used in the experimental phase, and of Mauro Cerrato, who contributed to the development of the platform.

**Author contributions** All the authors contributed equally to this work.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alem L, Tecchia F, Huang W (2011) Remote tele-assistance system for maintenance operators in mines. In: Underground coal operators' conference, pp 171–176
- Alexander T, Ripkens A, Westhoven M, Kleiber M, Pfendler C (2017) 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, pp 234–244
- Álvarez H, Lajas I, Larrañaga A, Amozarrain L, Barandiaran I (2019) Augmented reality system to guide operators in the setup of die cutters. *Int J Adv Manuf Tech* 103:1543–1553
- Aschenbrenner D, Latoschik M, Schilling K (2016) Industrial maintenance with augmented reality: two case studies. In: ACM Symposium on virtual reality software and technology, pp 341–342
- Aschenbrenner D, Rojkov M, Leutert F, Verlinden J, Lukosch S, Erich M, Latoschik M, Schilling K (2018) Comparing different augmented reality support applications for cooperative repair of an industrial robot. In: IEEE International symposium on mixed and augmented reality, pp 69–74
- Atheer, Inc ([acc. 23 June 2020]) Atheer AiR. <https://atheerair.com/>
- Battegazzorre E, Calandra D, Strada F, Bottino AG, Lamberti F (2020) Evaluating the suitability of several AR devices and tools for industrial applications. In: International conference on augmented reality, virtual reality and computer graphics, pp 1–20
- Besbes B, Collette SN, Tamaazousti M, Bourgeois S, Gay-Bellile V (2012) An interactive augmented reality system: A prototype for industrial maintenance training applications. In: Symp. Mixed & Augm. Real., pp 269–270
- Botto C, Cannavó A, Cappuccio D, Morat G, Nematollahi Sarvestani A, Ricci P, Demarchi V, Saturnino A (2020) Augmented reality for the manufacturing industry: the case of an assembly assistant. In: IEEE Workshop on everyday virtual reality, pp 1–6
- Cai Y, Wang Y, Burnett M (2020) Using augmented reality to build digital twin for reconfigurable additive manufacturing system. *J Manuf Syst* 56:598–604
- Chen CW, Chen WZ, Peng JW, Cheng BX, Pan TY, Kuo HC (2017) 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, pp 127–132
- Egger J, Masood T (2020) Augmented reality in support of intelligent manufacturing—a systematic literature review. *Comput Ind Eng* 140:106195
- Evangelista A, Ardito L, Boccaccio A, Fiorentino M, Petruzzelli AM, Uva AE (2020) Unveiling the technological trends of augmented reality: a patent analysis. *Comput. Ind.* 118:1–15
- Fang D, Xu H, Yang X, Bian M (2019) An augmented reality-based method for remote collaborative real-time assistance: from a system perspective. *Mob Netw Appl* 25:1–14
- Ferrise F, Caruso G, Bordegoni M (2013) Multimodal training and tele-assistance systems for the maintenance of industrial products. *Virtual Phys Prototyp* 8:113–126
- Fieldbit, Ltd ([accessed 23 June 2020]) Fieldbit Hero. <https://www.fieldbit.net/products/fieldbit-hero/>
- Fraga-Lamas P, Fernández-Caramés T, Blanco-Novoa O, Vilar-Montesinos M (2018) A review on industrial augmented reality systems for the industry 4.0 shipyard. *IEEE Access* 6:13358–13375
- Gattullo M, Scurati GW, Fiorentino M, Uva AE, Ferrise F, Bordegoni M (2019) Towards augmented reality manuals for industry 4.0: a methodology. *Robot Comput Integr Manuf* 56:276–286
- Guenther S, Kratz S, Avrahami D, Mühlhäuser M (2018) Exploring audio, visual, and tactile cues for synchronous remote assistance. In: Pervasive technologies related to assistive environments conference, pp 339–344
- Icona Srl ([accessed 23 June 2020]) Acty. <https://www.acty.com/>
- Jakl A, Schöffler L, Husinsky M, Wagner M (2018) AR for industry 4.0: architecture and user experience. In: Forum media technology, pp 38–42
- JoinPad ([accessed 23 June 2020]) Smart Assistance. <https://www.joinpad.net/smart-assistance-2/>
- Kim C, Chiu P, Tjahjadi Y (2019) A web-based remote assistance system with gravity-aware 3D hand gesture visualization. In: ACM International conference on interactive surfaces and spaces, pp 319–322
- Lamberti F, Manuri F, Sanna A, Paravati G, Pezzolla P, Montuschi P (2014) Challenges, opportunities, and future trends of emerging techniques for augmented reality-based maintenance. *IEEE Trans Emerg Topics Comput* 2:411–421
- Lang S, Kota MS, Weigert D, Behrendt F (2019) Mixed reality in production and logistics: Discussing the application potentials of microsoft hololens™. *Procedia Comput Sci* 149:118–129
- Liu C, Cao S, Tse W, Xu X (2017) Augmented reality-assisted intelligent window for cyber-physical machine tools. *J Man Syst* 44:280–286
- Masood T, Egger J (2020) Adopting augmented reality in the age of industrial digitalisation. *Comput Ind* 115:1–14
- MAXST Co, Ltd ([accessed 23 June 2020]) Vivar. <https://vivar.me/>
- Microsoft ([accessed 23 June 2020]) Dynamics 365. <https://dynamics.microsoft.com/>
- Mourtzis D, Zogopoulos V, Vlachou K (2017) Augmented reality application to support remote maintenance as a service in the robotics industry. *Procedia CIRP* 63:46–51

31. Mourtzis D, Zogopoulos V, Xanthi F (2019) Augmented reality application to support the assembly of highly customized products and to adapt to production re-scheduling. *Int J Adv Manuf Technol* 105:3899–3910
32. Muñoz A, Mahiques X, Solanes EJ, Martí A, Gracia L, Tornero J (2019) Mixed reality-based user interface for quality control inspection of car body surfaces. *J Manuf Syst* 53:75–92
33. Normand JM, Servièeres M, Moreau G (2012) A new typology of augmented reality applications. In: *Augmented human int. conference*, pp 1–8
34. Orlando O, Markon S (2017) The use of augmented reality in the maintenance of mechanical objects. In: *International conference on applied system innovation*, pp 842–845
35. Perla R, Gupta G, Hebbalaguppe R, Hassan E (2016) *InspectAR: An augmented reality inspection framework for industry*. In: *IEEE International symposium on mixed and augmented reality*, pp 355–356
36. Quint F, Loch F, Bertram P (2017) The challenge of introducing AR in industry: Results of a participative process involving maintenance engineers. *Procedia Manufacturing* 11:1319–1323
37. Re GM, Bordegoni M (2014) An augmented reality framework for supporting and monitoring operators during maintenance tasks. In: *Int. conference on virtual, augmented and mixed reality*, pp 443–454
38. RE'FLEKT ([accessed 23 June 2020]) RE'FLEKT Remote. <https://www.re-flekt.com/reflekt-remote>
39. Rice M, Chia S, Tay H, Wan M, Li L, Ng J, Lim H (2016) Exploring the use of visual annotations in a remote assistance platform. In: *ACM CHI conference on human factors in computing systems*, pp 1295–1300
40. Rice M, Ma KT, Tay HH, Kaliappan J, Koh WL, Tan WP, Ng J (2018) Evaluating an augmented remote assistance platform to support industrial applications. In: *IEEE world forum on internet of things*, pp 592–597
41. Sakata N, Kurata T, Kuzuoka H (2006) Visual assist with a laser pointer and wearable display for remote collaboration. *Trans Virtual Real Soc Japan* 11:561–568
42. Sanna A, Manuri F, Lamberti F, Paravati G, Pezzolla P (2015) Using handheld devices to support augmented reality-based maintenance and assembly tasks. In: *Int. conference on consumer electronics*, pp 178–179
43. Scope AR ([acc. 23 June 2020]) AR remote assistance. <https://www.scopear.com/solutions/ar-remote-assistance/>
44. de Souza Cardoso LF, Martins Queiroz Mariano FC, Zorzal ER (2020) A survey of industrial augmented reality. *Comp Ind Eng* 139:106159
45. Stream, Inc ([acc. 23 June 2020]) Stream™. <https://www.streem.pro/>
46. TeamViewer ([accessed 23 June 2020]) TeamViewer Pilot. <https://www.teamviewer.com/en/solutions/remote-support>
47. Techsee ([accessed 23 June 2020]) Techsee Live Contact Centers. <https://techsee.me/techsee-live-call-center/>
48. Techsee ([accessed 23 June 2020]) Techsee Live Field Services. <https://techsee.me/techsee-live-field-services/>
49. Vuforia ([accessed 23 June 2020]) Vuforia Chalk AR. <https://chalk.vuforia.com/>
50. Wang X, Dunston PS (2011) Comparative effectiveness of mixed reality-based virtual environments in collaborative design. *IEEE Trans Syst Man Cyber Part C Appl Rev* 41:284–296
51. Wang Z, Ong S, Nee A (2013) Augmented reality aided interactive manual assembly design. *Int J Adv Manuf Technol* 69:1311–1321
52. XMReality ([accessed 23 June 2020]) XMReality Remote Guidance. <https://xmreality.com/>
53. XOi ([acc. 23 June 2020]) XOi Vision. <http://www.xoi.io/solution/>
54. Zenati N, Bellarbi A, Benbelkacem S, Belhocine M (2014) Augmented reality system based on hand gestures for remote maintenance. In: *International conference on multimedia computing and systems*, pp 5–8
55. Zhu J, Ong SK, Nee AY (2013) An authorable context-aware augmented reality system to assist the maintenance technicians. *Int J Adv Manuf Technol* 66:1699–1714
56. Zubizarreta J, Aguinaga I, Amundarain A (2019) A framework for augmented reality guidance in industry. *Int J Adv Manuf Technol* 102:4095–4108

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.