Advanced Engineering Informatics 45 (2020) 101096

Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Structured authoring for AR-based communication to enhance efficiency in remote diagnosis for complex equipment

Iñigo Fernández del Amo[a,*], John Erkoyuncu[a], Rok Vrabič[b], Romain Frayssinet[a], Cristina Vazquez Reynel[a], Rajkumar Roy[c]

[a] Cranfield Manufacturing, School of Aerospace, Transport and Manufacturing, Cranfield University, United Kingdom
[b] Faculty of Mechanical Engineering, University of Ljubljana, Slovenia
[c] School of Mathematics, Computer Science and Engineering, City, University of London, United Kingdom

A R T I C L E I N F O

Keywords:
Augmented reality
Maintenance
Diagnosis
Remote collaboration

A B S T R A C T

Remote diagnosis procedures are prone to communication errors due to varying levels of experience and knowledge between expert maintainers and technicians. These result in inefficiencies that delay the diagnosis process. The aim of the paper is to develop a Structured-Message Authoring framework for Augmented Reality (AR) Remote Communication (SMAARRC) and to evaluate its ability to enhance the efficiency of remote diagnosis services. The framework proposes a message structure and automatic AR content creation rules for it that enable data capture and sharing within a remote context. Laboratory experiments present an average time reduction of 56% for remote calls while maintaining same quality compared to traditional remote communication methods (phone calls and emails). Remote experts feedback evidence the usability and feasibility of this framework to work in real-life conditions.

1. Introduction

Engineering collaboration is a socially-mediated technical activity that involves multiple people working interdependently to achieve a greater goal than is possible for any individual to reach alone [1]. The progress in information-communication technology has made it possible for the collaboration to take place remotely over large distances, allowing globally dispersed businesses to operate around the clock. Many engineering processes, such as collaborative design [2] and remote maintenance [3], have been improving using remote interfaces. Remote engineering collaboration is increasingly becoming necessary in remote maintenance for various reasons such as reducing cost of travel and increasing efficient use of experts' time. A recent report on UK service and support industry valued the global market in 'service and support' across high value manufacturing sectors at £490 billion in 2017 [4]. The remote maintenance market is poised to grow at around 25.9% over the next decade to reach approximately $69.81 billion by 2025 [5].

Remote diagnosis is one of the biggest challenges in remote maintenance of complex equipment [5] (e.g. machine tools or aerospace machinery). Remote diagnosis refers to remote support involving procedures for finding failures, and validating and fixing components' faults that contribute to them. It is usually conducted with the guidance of expert remote maintainers and implemented by on-site technicians. The technicians can have varying levels of experience and knowledge, and require guidance from the remote expert. In these situations, enabling efficient communication is essential. Communication challenges such as task misunderstanding or component misplacement can cause delays and errors, and ad-hoc communication (e.g. phone calls) may lead to more confusion. Hence, communication should be structured in order to be efficient.

A very effective interface in remote maintenance contexts is Augmented Reality (AR). Compared to conventional approaches (e.g. phone calls, emails), AR can enhance communication efficiency by its ability to overlay virtual information into real-life scenarios [6]. For example, it can enable the expert to point in technician's sight the exact location to which an instruction is referred to. One of the main advantages of using an AR application is that complex information can be represented in a more comprehensible format, making it easier for the operator to understand [7]. AR studies in maintenance show promising results for enhancing human performance when carrying out maintenance tasks [8] and show that AR can significantly improve maintenance efficiency in comparison to existing information-delivery methods (e.g. paper manuals, phone calls, etc.) [9]. However, several

* Corresponding author at: Building 30, Cranfield University, Cranfield MK43 0AL, United Kingdom.
E-mail address: i.fernandez@cranfield.ac.uk (I. Fernández del Amo).

https://doi.org/10.1016/j.aei.2020.101096
Received 16 August 2019; Received in revised form 27 March 2020; Accepted 30 March 2020

areas (e.g. authoring, context awareness, and interaction analysis) still remain challenging for efficient use of AR to support maintenance tasks [10]. A general concern within those is to identify the information that can be displayed for effective AR support of a maintenance task [5]. In AR for remote maintenance particularly, there is lack of research to characterise the language and the structure of communication for efficient information delivery [7]. Since no universal or standard language exists, a common approach is to define a set of agreed-upon symbols and then leave some space for free input.

This paper goes a step further and proposes a framework with a common message structure for AR-based communication in the remote diagnosis context. It tackles the challenges related to ambiguity and inaccuracy of communicating tasks and associated components by developing an AR-based solution for accurate and timely remote guidance. This research makes two major contributions to the AR and remote engineering collaboration literature in the context of improving diagnosis efficiency:

- It develops a structure that determines elements of message codification to minimise the above-mentioned remote communication challenges.
- It proposes a new rule-based approach to automatically generate AR content from structured messages for efficient communication in remote diagnosis processes.

2. Literature review

2.1. Remote collaboration: challenges and opportunities

Remote collaboration can play an important role in maintenance [11]. As an example, it is emphasised that equipment repairs and maintenance cost between 30% and 50% of the total operating costs of mine sites. Moreover, studies [12] also predicted that “every 1% improvement in equipment availability or productivity improves the overall productivity by up to 3.5%”. Remote collaboration can support such improvements by offering efficiency and effectiveness in delivering maintenance tasks [13].

In face-to-face collaboration, participants commonly have a shared understanding of what is happening in the present workspace. Along these lines, technologies that can offer a range of information about what a collaborator is currently doing or intends to do shortly are a key challenge in remote collaboration. In addition to limited technological capabilities, requisition of suitable lines of communication and lack of internal standards are just a few symptoms of ineffective remote collaboration [14].

Distance promotes the need for remote collaboration. The challenge with remote collaboration is not only related to distance that is driven by physical attributes, but also operational and affinitive. Affinitive distance involves the resultant gap in operational styles between team members, and its effect on synergy, camaraderie, and management. On the other hand, operational distance refers to that of communication between teams of various sizes, as well as the gap between the skill levels of different team members [15]. All of these factors result in affecting the efficiency of remote collaboration.

Remote diagnosis scenarios mostly involve “collaborative physical tasks”. These are tasks in which various individuals work alongside to conduct activities in real-world objects [16]. Several observational studies [17–19] suggest that communication in collaborative physical tasks is directly related to target object identification, activities description, and successful performance confirmation. These elements can be described as the “situational awareness” [16]. Besides, it has also been pointed out [20] that situational awareness in remote collaboration is managed through the use of visual information. Hence, it can be said that the higher the visualisation of information, the higher the situational awareness; and so, the higher efficiency of remote collaboration.

A technology that can enable information visualisation for enhancing efficiency in remote collaboration is AR. The introduction of AR in remote collaboration processes, particularly for knowledge intensive works, is important for numerous reasons [21]:

1. reducing the costs of maintenance tasks
2. reducing the risk of accidents that may occur
3. improving time taken to complete a task
4. improving the communication between technicians and experts
5. reducing experts’ expenses for traveling to remote sites

The ability to overlay spatially meaningful information on the 3D space allows the AR technology to be a promising option to support knowledge-intensive work [22].

2.2. Uses of AR in remote collaboration

Remote collaboration is an emerging trend in AR research [23], especially in the area of maintenance [24]. Publications in the area of AR remote maintenance can be classified based on the nature of the maintenance tasks being supported. The nature of maintenance tasks researched in AR remote maintenance is mainly physical. Although from a maintenance view, these can be of different procedural nature such as repair, assembly or diagnosis [25].

AR remote collaboration in assembly tasks seemed to have attracted lots of research interests. Gurevich et al. [26], Fox [27] and Adcock and Gunn [28] presented some relevant examples. These proposals focused on providing remote support to enhance situational awareness of components to be assembled and/or disassembled by allowing remote experts to interact with real components through 3D model replicas. Other interesting examples are those from Oda et al. [29] and Zenati et al. [30]. These enable the hands of the remote expert to be tracked and rendered right in front of the technician’s view. So, the expert can help the technician to navigate towards the correct object to focus on.

Similar approaches are proposed by those works focused on AR remote collaboration for maintenance repair. One example is that of Rambach et al. [31]. They propose a tracking system that enables additional objectual awareness support by enabling remote experts to share textual notes and other holograms (e.g. arrows, circles, etc.) than component replicas completely aligned with the on-site technician’s point of view. That is also the case of the proposal from Reddy et al. [32], which enables to create 3D models of components using single images.

Besides applications in assembly and repair, there also exist proposals in AR remote collaboration for diagnosis and/or inspection. These cases require of additional situational awareness support in terms of the procedures to conduct. So, the AR systems proposed provide further support through step-by-step guidance. One relevant proposal is that from Mourtzis et al. [33]. They proposed a system where the remote support is triggered by equipment control monitoring using a product-service system approach. When the control monitoring system advises for a preventive maintenance routine, the expert is connected with an on-site technician to first inspect the condition of the equipment and then support the repair procedures when necessary. These tasks are supported through AR visualisations that include 3D models and free-text messages from the remote expert. A similar approach is also suggested by El Ammari and Hammad [34] and Negges et al. [35]. The former utilises BIM technologies to share the content between a remote expert using VR (Virtual Reality) and an on-site technician using AR and to enable. The later also implements GPS indoor tracking technology to improve technician’s navigation while inspecting complete buildings. On a separate note, Hadar et al. [36] suggest providing preventive support by means of Artificial Intelligence (AI) in order to reduce the work overload of remote experts. Their proposal consists of an AI “bot” that guides on-site technicians through initial procedural routines using AR content. If the “bot” is not capable of helping the technician to resolve the issue, then the remote expert can connect to provide additional, more expert support.

Research in AR for remote maintenance collaboration focused on developing different AR techniques (e.g. tracking and authoring) to provide support for various maintenance tasks (e.g. repair or diagnosis). These proposals consider enhancing collaboration by improving situational awareness of remote technicians utilising different AR content types (e.g. 3D holograms, textual notes, etc.).

2.3. AR authoring techniques in remote collaboration applications

AR authoring techniques comprise software methods, tools and algorithms to create augmented content for AR applications [10]. In the case of remote maintenance, authoring techniques refer to the means that remote experts have to transmit messages using different AR content types (e.g. holograms or textual notes) to on-site technicians and vice versa. These authoring techniques can be classified by two different aspects. The first one is the AR content types enabled in AR applications. The second aspect is the balance between the two directions of communication, both from the remote expert to the on-site technician and vice versa.

With respect to AR content types, different research works suggest different variations of them. There is some evidence of authoring techniques that only provide support through using virtual replicas of the remote experts hands [29,30]. Although these approaches enhance situational awareness of on-site technicians, they can be limited in situations where additional information is required (e.g. specific torque to screw a bolt). Other research works provide authoring tools for remote experts that enable the use of 3D holograms. In some cases, these 3D holograms are limited to virtual replicas of real-life components [27], In other cases, 3D holograms also include more specific forms (e.g. arrows, circles) [26,28] that help to further explain the details of a specific procedural step (e.g. direction of screwing). Nevertheless, most authoring techniques [31–35] enable the remote expert with a combination of AR content types (holograms and textual 3D notes). So, they can freely decide which one is best for each specific message to be sent. Although, this approach has the risk of cognitive overloading the on-site technician with lots of information. Thus, resulting in decrease of maintenance efficiency and safety.

With respect to balance of communication, not all research proposals show evidence of a communication protocol that ensures information is being sent and received, from expert to technician and vice versa. Thus, resulting on an effective collaboration. Some research works [32,34,35] provide a “chat box” that enables un-regulated two-way communication via text. Some others [27,30–32] do not enable communication from the on-site technician to the expert. Most of the papers analysed in Section 2.2 simply assume the remote collaboration occurs for guidance purposes and leave the remote expert the decision on when to send the next message. This could cause confusion in case the remote communication suffers from latency issues or in case the on-site technician cannot transmit issues on understanding the expert’s messages.

Overall, AR authoring techniques for remote collaboration enable to transmit messages with a variety of content types (e.g. 3D models and textual notes). Besides, they also implement one-way regulated or un-regulated two-way communication protocols. Nevertheless, there seems to be lack of evidence of authoring techniques establishing regulated two-way protocols for communications and correlations between the elements of messages being sent and the AR content types used to transmit them.

2.4. Research gaps

AR research in remote collaboration has focused in the following areas [37]:

- Spatial problem solving: to enhance the ability to recognise objects and their virtual counterparts.
- Cognitive interaction: to adapt content to the cognitive workload of the technician.
- Interactive design: to investigate performance, behavioural and cognitive effects of virtual contents.

These areas tend to put the technology at the centre of delivering effective remote collaboration. Nevertheless, there is lack of research evidence focusing at methods that ensure appropriate levels of situational awareness regarding the messages being sent between experts and technicians in remote communications and the protocols to do so. A possible method could consider the creation of message structures for communication exchange in remote collaboration to enhance situational awareness of technicians through a better mutual understanding with experts. That could also have further impacts in remote diagnosis research as it can allow to record, and reuse data exchanged in previous communications for future remote diagnosis operations or other maintenance areas (e.g. failure prediction). Hence, this research aim is to propose an AR solution that includes structured messages and communication protocols to enhance the diagnosis efficiency by improving the situational awareness of the remote collaboration.

3. Methodology

The research aim identified in the previous section drove the selection of an appropriate research method to successfully design and validate an AR research solution. Inspired by similar works in the field [6,37,38] and well-established methodologies for design research [39], the method utilised by this research is as follows:

1. Identification of objectives: conduct a literature review [40] to find specific research gaps and justify the value of the research solution proposed. The results were presented in Section 2.
2. Solution design: utilise the 5 W’s method [16,39,41] to define the structure of remote messages and the communication protocol. Then, map [42] the resultant message elements against relevant AR content types to enhance situational awareness of remote communication. The proposed solution is presented in Section 4.
3. Demonstration: conduct interviews [43] with experts in remote diagnosis to identify relevant scenarios and produce a case study for the research solution’s validation. Then, implement the proposed solution in a prototype AR system for further experimentation. The case study and the system implementation are presented in Section 5.
4. Validation: design an experiment [43] according to relevant criteria in remote collaboration [16,20] to validate the research proposed. Then, conduct the experiments, and analyse and assess the results. The experiment design is presented in Section 5. Section 6 discusses the experimental results and compares them with relevant literature. The final conclusions of this research, along with proposed future works are presented in Section 7.

4. Structured message authoring for AR-based remote communication (SMAARRC)

This research proposes an AR-based remote communication framework based on: (1) an innovative message structure (4.1), and (2) a rule-based authoring approach for automatic AR content creation (e.g. holograms, images, etc.) (4.2). Fig. 1 presents this approach compared to conventional remote collaboration methods. In conventional approaches, remote communication is enabled through methods such as emails or phone calls. Such methods allow for unstructured text or conversations to be sent. And so, it is more difficult to ensure that the messages convened by them are situational aware. Instead, this structured-message authoring AR approach arranges messages in different

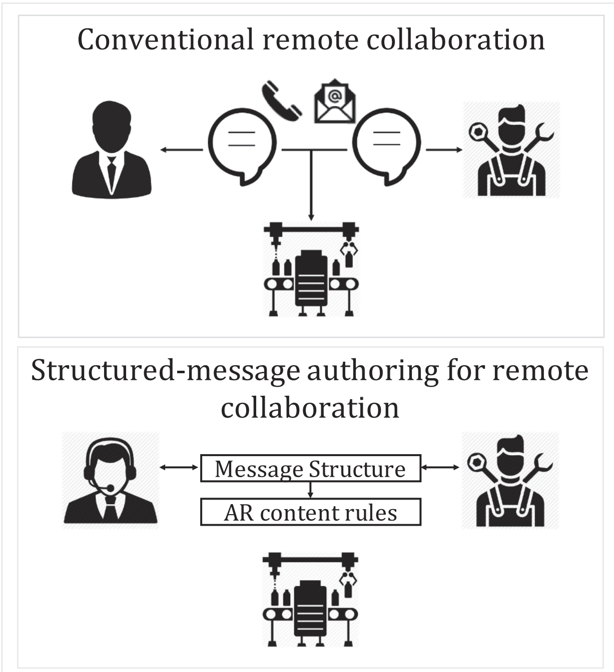

Fig. 1. Description of structured message authoring to enable situational awareness for enhancing efficiency in remote collaboration.

components to include situational awareness elements (e.g. component location). It then converts some of this message elements in AR content to enhance the efficiency of the communication (e.g. component highlighted in technician's field of view).

4.1. A message structure to comply with situational awareness needs of remote collaboration

A remote communication between an expert and a technician can be described as a '**call**', which involves an exchange of '**messages**'. A message, verbal or written, is made up of a set of **'message elements'** strictly adhering to a **'message structure'**. This offers a consistent language and structure for remote communication in the diagnosis context.

The Five W's method [41,43] was used to provide a mutually exclusive, collectively exhaustive set of message elements that complies with remote diagnosis challenges:

- Ability to declare messages that can describe any procedure.
- Ability to record and replay a call based on the message logs.
- Ability to use 'message elements' for creating AR content.

The Five W's method uses six basic questions: Who? What? Where? When? Why? How? It is used to develop a structure that sets a rule to codify messages for efficient communication. This method is employed to identify the '**message elements**':

- ***Who?*** Messages can be sent by one or more *experts* and *technicians* that act as '**senders**' or '**recipients**'.
- ***What?*** Messages in remote diagnosis should define the '**type**' of processes being described. In a diagnosis procedure, this '**type**' can be: *action*, *confirmation*, *question*, or *response*.
- ***Where?*** Messages should also indicate the '**place**' where the process is occurring. This can be described in the form of a '**component**' and/or a '**location**' related to it.
- ***When?*** Messages have a specific order within a call. A message '**identifier**' allows to reconstruct the order of the messages within a call.
- ***Why?*** Messages should clarify the context or '**category**' of the process being described within a diagnosis call. In line with the definition above, a '**category**' can be: failure *definition* or component's fault *validation*.
- ***How?*** Messages should describe the maintenance methods used to conduct the procedure being described. A method can be described in the form of an '**action**' and a '**measure**'.
- In case the previous elements cannot describe accurately a message, an '**object**' could be added for further clarification.

These message elements and the values they can take to generate messages are presented in Table 1. The proposed message structure focuses on diagnosis and considers the full spectrum of the remote communication requirements. It helps the diagnosis problem by ensuring the scope of diagnosis and associated sequence of steps is efficiently followed.

In order to send a message using this structure, a sender has to provide a value for all the message elements except **'Object'**, which is only used for further clarifications. Then, recipient(s) have to reply with a confirmation message before sending other messages. This establishes a communication protocol that determines what message is being shared by collaborators at any time. This helps to ensure that the AR content for each message is not overlapped with content from other messages in the augmented scene. Fig. 2 shows an example message and the AR contents for its message elements in both, technician's and expert's view. The technician's view comprises the augmented scene. The expert's view includes a virtual environment to interact with the equipment's model and code messages, and live-streaming of the technician's view to ensure correct message execution. This message structure can also help to analyse remote diagnosis tasks. It allows to store and process messages in a structured way. The following section describes the rules and AR content types to augment message elements.

Table 1
Remote diagnosis messages' structure: elements and values.

Element	Definition	Values
Sender	Person that sends a message	Expert or Technician
Recipient	Person(s) that receive(s) a message	Expert(s) or Technician(s)
Identifier	Order of a message in a call	Integer or date and time
Category	Diagnosis context of the message	Definition or Validation
Type	Aim of the message being sent	Action, Confirmation, Question, or Response
Component	Equipment's part a message refers to	Component's name in its CAD model
Location	Place to where a message refers to	Position and rotation from component
Action	Method to conduct a task being defined	Pull, Push, Screw, Inspect, Measure…
Measure	Magnitude for applying the method	Quantitative measure Qualitative measure
Object	Additional data to complete the message	Free text

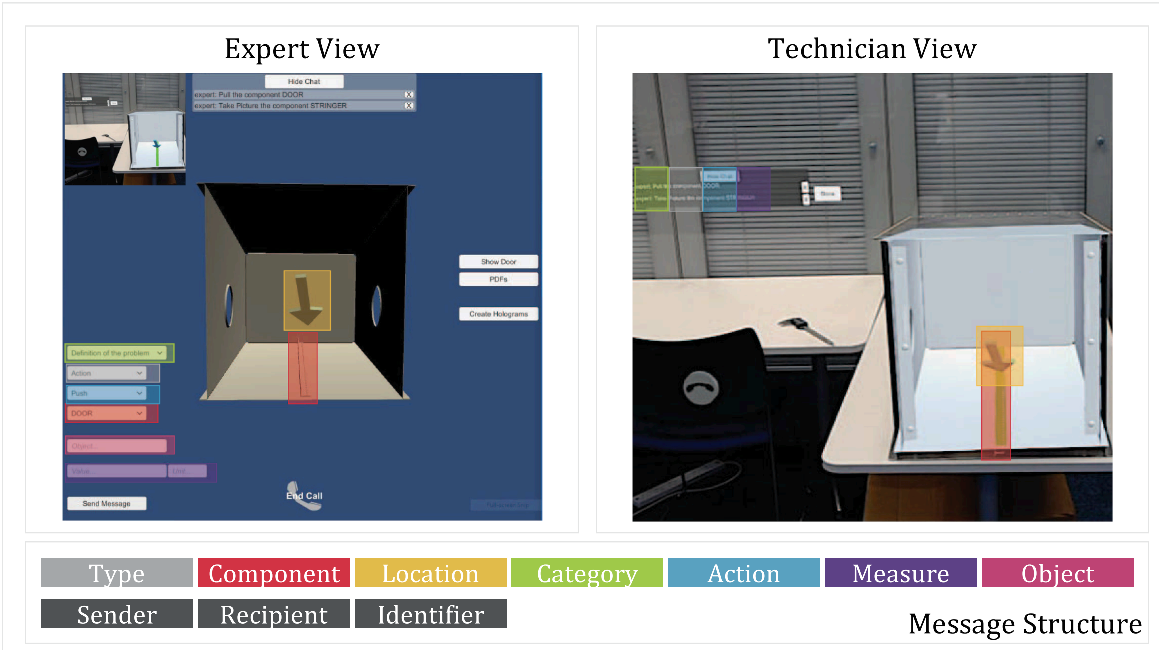

Fig. 2. Example of message structure and AR visualisation: expert vs technician.

4.2. Rules-based authoring for message elements automatic AR content creation to enhance diagnosis efficiency

Augmentation refers to the methods deployed to enhance natural environments or situations with virtual information in order to offer perceptually enriched experiences [10]. The augmentation methods used are Holograms (H), 3D Measurements (M), Textual values (T), and Pictures (P). These enhance visualisation of message elements to ensure situation awareness and enhance remote diagnosis efficiency. Fig. 3 presents few examples including all augmentation methods (AR content types).

Holograms are three-dimensional images used to overlay the real scene with digital artefacts such as arrows, circles or component's 3D models in order to mark a particular feature of the scene [8]. The expert can deploy holograms remotely to provide guidance to the local technician with the aim to increase object awareness. This is achieved by giving the expert the choice to allocate holograms with their preferred shape, position, scale, and rotation. These are then communicated through the message elements including 'measure', 'location', and 'component'.

3D Measurements are used to make the messages more accurate. They provide more precise values for the message elements 'measure' and 'action' in real-time by using Bluetooth devices or the so called bare-hand interaction [6] gestures for metric measurements.

Textual values are used as an augmentation method for multiple purposes. For example, it overlays the predefined questions to derive diagnosis and it helps to ensure the recipient uses the same vocabulary as the sender. This increases procedure awareness for message elements including 'location', 'component', 'action', and 'measurement'.

Pictures can provide additional details to increase procedure and object awareness for 'location' and 'object' message elements. For example, when technicians send real-time pictures to experts for evaluating the condition of a component's surface.

Mapping 'message elements' against these augmentation methods (AR content types) identifies the rules for automatic authoring (content creation). The mapping process consisted of identifying 'message elements' relevant for situational awareness and determining visual AR content types that can enhance their visualisation for faster message understanding. Collaboration situational awareness can be divided in

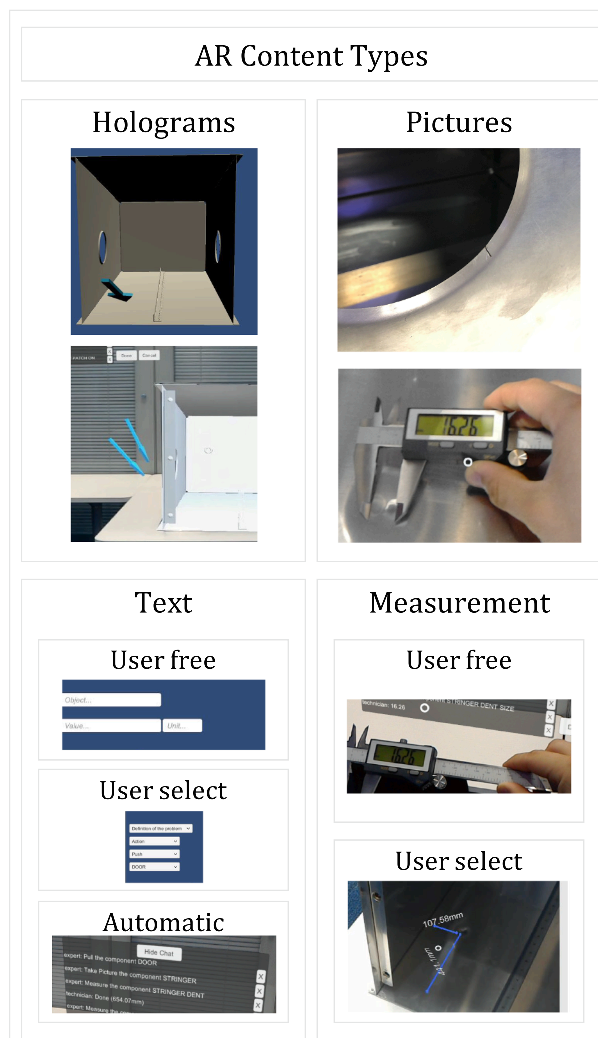

Fig. 3. Examples of augmentation methods implemented in SMAARRC.

Fig. 4. Examples of content-creation rules for a message being sent from the remote expert to the technician.

object and procedure awareness. Object awareness refers to the ability of messages to identify the real-world objects being referred [20]. Instead, procedure awareness refers to the ability of messages to clearly define the task being referred [16]. Depending on whether a ‘message element’ can affect any of these two aspects, an augmentation method is assigned through an authoring rule. The rule-based authoring method therefore determines the content to be overlaid on the expert and technician views (Fig. 5), when a message is to be received or created while remotely collaborating. An application example of these rules concerning a message being sent from the remote expert to the on-site technician is given in Fig. 4. Fig. 5 presents the same example from both views, technician and expert (which can include live-streaming of technician’s view).

These rules regulate how to augment message elements using AR. Each rule identifies the pre-determined augmentation method for each message element. For example, Holograms (H) are used to overlay the ‘**component**’ message element on technicians’ field of view for increasing situational awareness of components being referred by messages. The complete set of rules (object and procedure awareness) and their different applications to the remote expert and on-site technician views are presented in Fig. 6.

5. Validation protocol

5.1. System implementation

The proposed solution was implemented within a prototype AR system for experimentation. This system comprises three components: (1) a desktop computer with a virtual interface for the remote expert (including technician’s live-streaming), (2) a cloud server to store the transferred data, and (3) an HMD (Head-Mounted Device) for the technician with an AR application using HoloLens. The system architecture, along with the software and hardware used to build it, is presented in Fig. 7.

5.2. Experiment design

The validation experiments aim at evaluating the ability of the proposed solution to enhance efficiency of remote diagnosis operations by improving the situational awareness of remote messages. Their objectives are to collect data for demonstrating the following hypothesis:

- There is no significant difference in errors results by solution (AR vs NOAR).
- There is a significant difference in time results by solution (AR vs NOAR).
- Correlation between solution and object awareness effects is significant on time results.
- Correlation between solution and procedure awareness effects is significant on time results.
- Real-life experts and testers consider the solution proposed useful to enhance efficiency of remote collaboration in diagnosis scenarios.
- Real-life experts consider the solution proposed feasible to be implemented in real-life conditions on remote diagnosis scenarios.

According to similar research [6,16,37], the following measures can be considered as appropriate for such evaluation:

- **Quantitative:** time and errors of remote diagnosis operations.
- **Qualitative:** usability and feasibility of AR methods supporting remote diagnosis tasks.

For these measures to be appropriate for evaluating efficiency enhancements, the following assumptions must hold true:

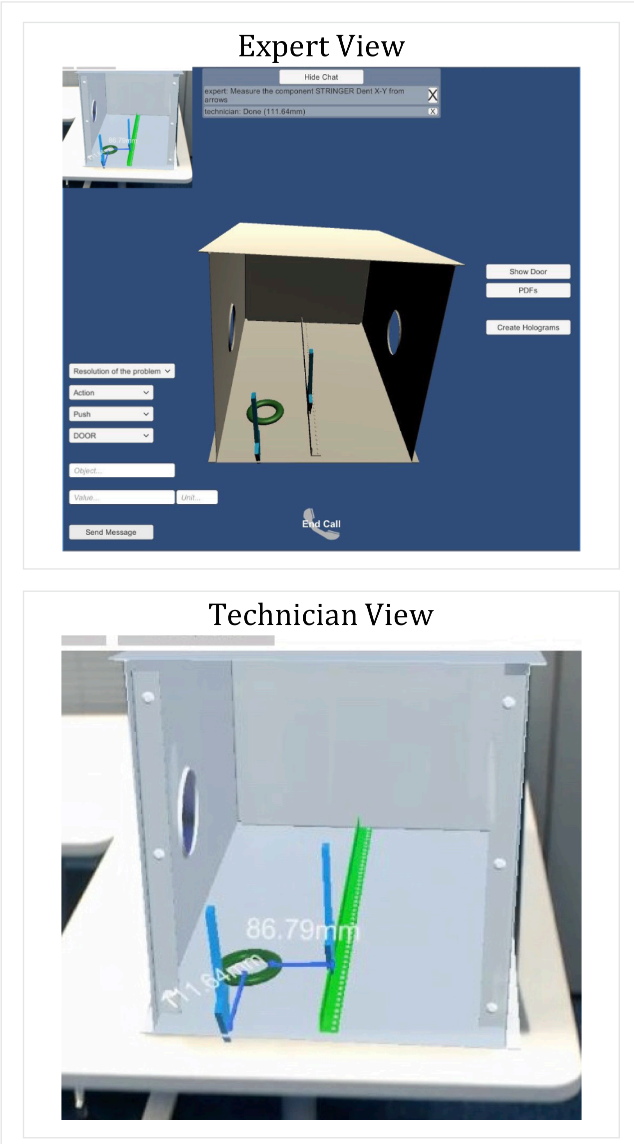

Fig. 5. Example of AR content creation: A message from the expert and the technician views.

- Time and errors can be a direct representation of efficiency if consistent quality of remote diagnosis operations is assumed. In order to ensure that, the study assumes a pre-determined operation, the quality of which does not depend on the technician performance.
- The usage of the AR solution can affect the efficiency in the cases where it is not compatible with technician's manual operations (e.g. hands use), the working environment (e.g. light conditions) or the technician does not know how to use it properly. So, its feasibility and usability are important measures towards the evaluation of the quality of a remote diagnosis operation. Since there are no available quantitative measures to evaluate those, qualitative, subjective measures have been chosen.

Quantitative and qualitative measures involve different research methods to be collected and their results analysed. These are presented in the following subsections.

5.2.1. Stopwatch time and errors study

The stopwatch time and errors study aims at analysing the effect of the proposed solution on situational awareness within remote collaborative diagnosis compared to alternative solutions (e.g. phone calls or emails). It is assumed that enhanced situational awareness of remote messages increases efficiency of diagnosis tasks. In such scenarios, it can be said that efficiency depends only on time for similar levels of effectiveness (quality).

Time can be defined as the number of seconds required by a technician to receive a message from a remote expert and complete the associated tasks. Quality, also declared as **errors**, is defined as the number of tasks completed by a tester which may involve a deviation in form or result of what is declared by the corresponding remote message. Besides, situational awareness refers to the ability of remote messages to explain the state of real-world objects (object) and the tasks to conduct (procedure). These two classify the different levels of complexity of a remote message.

Based on previous definitions, it can be said that if errors (quality) are invariable, then the effect of AR solutions on situational awareness can be evaluated based on its effect on completion time. In order to test that, an experiment (study) can be designed as follows [43]:

1. Declare a remote diagnosis operation that includes common operations in the maintenance context.
2. Declare remote messages that cover all operational steps and include all levels of complexity of situational awareness.
3. Conduct experiments with those messages to study the effect of alternative solutions (AR and non-AR) on sending those messages by measuring **time** and **errors**.

Section 5.3 describes the remote diagnosis operation, messages and equipment used to conduct these experiments. These elements were designed with the support of real-life experts, who helped to identify the necessary requirements to match real-life working scenarios.

If the assumptions above are correct, then it is reasonable to expect the following results:

- Errors do not vary with the use of AR or non-AR solutions.
- Completion times are reduced with the use of AR solutions compared to non-AR solutions.
- Differences between completion times for AR and non-AR solutions increase when the level of complexity of situational awareness increases.

The study described above considers one variable to test assumptions (**errors**), a response variable (**time**), and three independent factor variables (**AR usage**, **object awareness** and **procedure awareness**). While the two variables have been defined above, the factor variables are defined in Table 2.

Each factor variable has different levels. Their definitions are presented in Table 3.

The experiment aims at testing all levels of complexity of situational awareness. In order to do so, it is necessary to declare a remote collaborative diagnosis operation that includes remote messages with all those levels. These messages and their situational awareness complexity levels are presented in Section 5.3. Each experiment of the study will consist of a tester conducting the diagnosis tasks related to each message received. Time will be measured for each message, while errors will be measured for each tester. Errors measurements are taken to validate the assumption of maintained quality among solutions.

5.2.2. Usability and feasibility questionnaires

The usability and feasibility questionnaires aim at evaluating the perceived validity of the AR methods utilised to deliver information for remote diagnosis support. Usability and feasibility are criteria that refer, respectively, to the ability of the AR solution to deliver information in an appropriate manner from the expert to the technician from user and working environment perspectives. These criteria have been divided in different sub-criterions to cover independently each

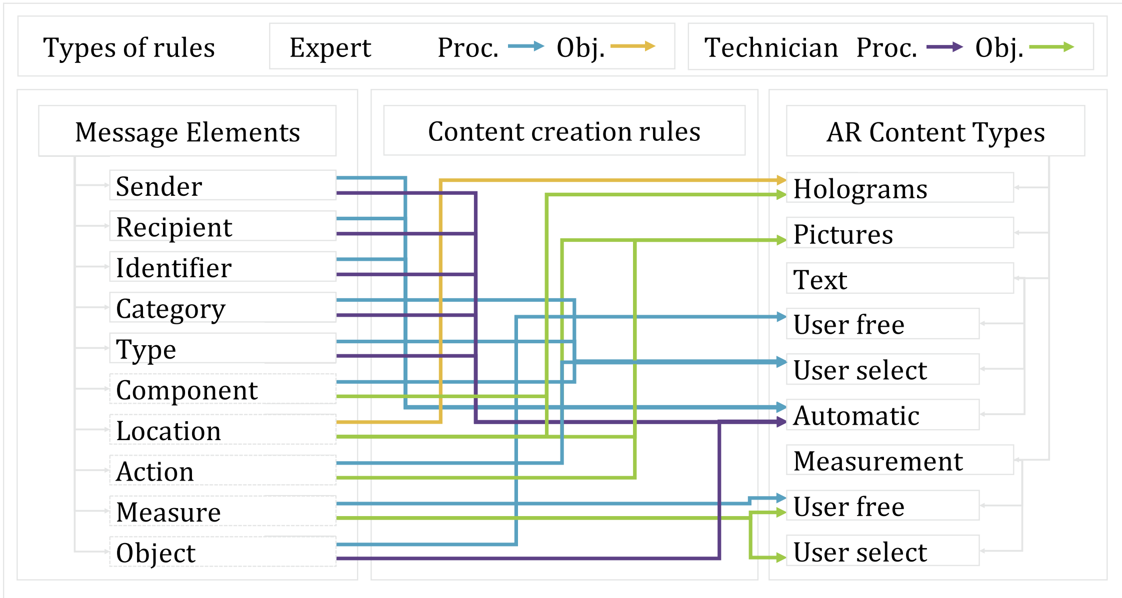

Fig. 6. Descriptive schema of all content-creation rules based on view and object and procedure awareness.

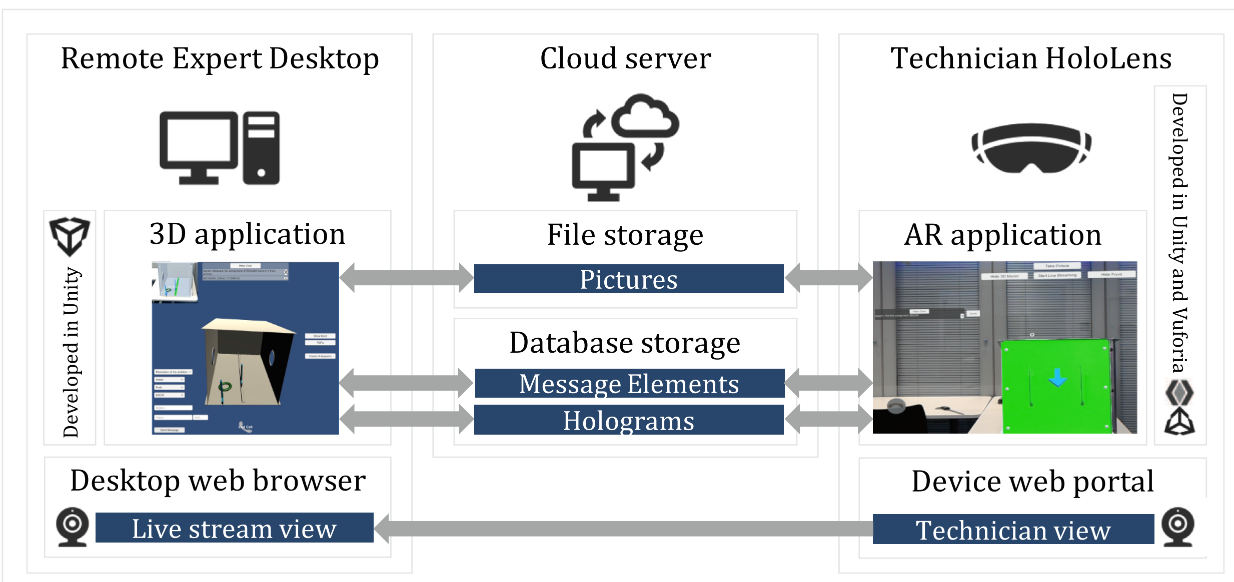

Fig. 7. Description of the AR-based system implementation.

aspect of the AR methods used as well as the operation quality discussed above. The criterions utilised are based on Nielsen's usability criteria [44] already utilised in similar research [6,45–47]. These criteria, its evaluation elements and methods are presented in Table 4.

Each evaluation element is presented as a question according to the scale defined in two independent questionnaires for each main criterion. The results collected are utilised to evaluate the usability and feasibility of the proposed AR methods in a remote diagnosis environment according to the case study proposed below. Besides, operational quality is also evaluated by potential time savings. Some assumptions are considered:

- Errors are not evaluated in quality terms as they may be dependent on user expertise, which can vary.
- It is assumed that the quality is of consistent level for the stopwatch time study if the results of the questionnaire provide a similar result to the experiments.

Table 2
Definition of relevant factors in stopwatch time and errors study.

Factor	Definition
AR usage	Utilisation of an AR solution to send and receive messages and confirm completion of associated tasks
Object Awareness	Ability of a message to identify the real-world object to which tasks refer
Procedure Awareness	Ability of a message to identify the steps to successfully conduct the activity referred

Table 3
Definition of factors levels of stopwatch time and errors study.

Factor	Level	Definition
AR usage	AR	Utilisation of the proposed AR solution by expert and technician
	NOAR	Utilisation of phone calls and emails for remote communication by expert and technician
Object Awareness	Simple	Message includes words that explicitly declare the object(s) subject of the task
	Complex	Message does not include words that explicitly declare the object(s) subject of the task
Procedure Awareness	Simple	Message includes words that explicitly declare the action(s) to conduct the task to satisfaction
	Complex	Message does not include words that explicitly declare the action(s) to conduct the task to satisfaction

Table 4
Description of criterions, evaluation elements and methods for usability and feasibility surveys.

Criteria	Sub-criteria	Element	Evaluation
Usability		Ease-of-use	Likert Scale (1–5)
		Ease-to-learn	Likert Scale (1–5)
		Satisfaction	Likert Scale (1–5)
	Utility	User interface	Likert Scale (1–5)
		Button interaction	Likert Scale (1–5)
		Picture interaction	Likert Scale (1–5)
		Message log	Likert Scale (1–5)
	Quality	Time improvements	Likert Scale (1–5)
Feasibility		User interface	Binary Scale (yes–no)
		Button interaction	Binary Scale (yes–no)
		Picture interaction	Binary Scale (yes–no)
		Message log	Binary Scale (yes–no)
		Data usage	Binary Scale (yes–no)

The methods to collect data according to the tests and surveys above are explained in the following subsection.

5.2.3. Experimental and analysis protocol

The experimental protocol describes the necessary steps to collect and analyse relevant data regarding the abovementioned qualitative and quantitative variables for validating this research's proposal. This protocol comprises four steps, two for data collection and two for data analysis. The protocol for data collection is as follows:

1. **Real-life expert semi-structured interviews:**
a. **Solution demonstration:** to allow real-life experts to utilise both ends of the proposed solution and collect their opinions on real-life testing scenarios (equipment and operations).
b. **Post-demonstration expert questionnaire:** to capture qualitative data of expert feedback on the proposed solution's usability and feasibility.
2. **Laboratory experimentation:**
c. **Stopwatch time and errors study:** to capture quantitative data on the effect on time and errors of the proposed AR solution compared to existing alternatives (NOAR: phone call and email) according to various complexity levels of remote messages.
d. **Post-experimental tester questionnaire:** to capture qualitative data of tester feedback on solution's usability.

The reasons to separate expert interviews from experiments were two. First, experts' opinions were necessary to design experiments based on real-life operations. Second, the experiments sample size (as calculated in Section 5.3) was bigger than the resources available for this research in terms of real-life technicians and working environments. That is why the protocol declared above utilises university students as testers and real-life experts. While the experiments aim to compare the proposed solution with currently existing alternatives, the questionnaire responses from testers and experts can be compared to analyse the usability feasibility of the proposed solution.

This protocol considers some assumptions:

- Testers are not questioned on solution's feasibility. This is because their inability to provide reasonably valid feedback as they are not subject matter experts.
- Feasibility is measured in a binary scale as there is no option for results comparison. So, the research question is whether the AR methods are feasible or not.
- Usability is measure on a Likert scale in order to compare results different interviewees according to their expertise in the subject. If experts and testers provide similar responses, then it can be said that content is usable in remote diagnosis environmental conditions.

The experimental protocol also includes the analysis of data collected. That is as follows:

1. **Errors effect:** to ensure that "quality is maintained among experiments" is a valid assumption. Basic statistics and graphical evaluations will be utilised for this matter.
2. **Effect on time:** to study the effect on time of the proposed solution according to message complexity factors compared to current alternatives. The response variable of the experiment is completion time. Factor replication tables show the number of measures taken on the response variable for each tester. Regarding the assumptions of the experiment design, it is important to analyse the correlation between the effects of the response variable with the different factors. In order to do so, a three-way analysis of variance (ANOVA) has been chosen as the statistical approach. This includes post hoc test comparisons for the interactions in the form of a Tukey HSD test. The results of such statistical analysis can only be considered valid if the tests assumptions are correct. Nevertheless, it will also be convenient to evaluate the assumptions of homogeneity, normality, linearity, and additivity.
3. **Questionnaire results:** to study the different criteria utilise to validate from a user perspective the usability and feasibility of the solution proposed. Results quantify qualitative responses. Hence, basic statistics and graphical evaluations will be used to analyse them.

As a result of the expert interviews, the protocol above was applied to a case study from the aerospace industry. Hence, this research conclusions should be discussed within that context. The case study is presented in the subsection below to provide the necessary context for the analysis conducted in following sections.

5.3. Case study

The case study comprises a remote diagnosis operation, which, in turn, comprises the messages to be sent from experts to technicians, and the equipment with which the operation is conducted. The case has been selected based on the data given by diagnosis experts from an aerospace company. The company provided over 20 h of interviews (Section 5.2.3) to identify current processes and solutions in remote diagnosis, which were the basis for the design of the case study.

The remote diagnosis operation is a combination of different remote-driven tasks which are common in the context of the company (e.g. visual inspection, photograph, repair, etc.). These tasks can be

Table 5
Description of messages that comprise the case study's remote diagnosis operation.

Message	Description
A	Expert asks to unscrew the screws of the front panel of the fuel hatch and open it
B	Expert asks to visually inspect the right and left sides of the hatch and to take a photograph of every defect found
C	[Two defects should be found by tester] Expert asks to repair by placing the patch
D	Expert asks to take a photograph of the previous reparation result and send it by email

Table 6
Description of validation criteria (OA and PA) factor levels of case study messages.

Message	OA	Reasoning	PA	Reasoning
A	Simple	Action objects declared	Simple	Actions declared
B	Complex	Orientation of action objects not declared	Complex	Action not declared to enough quality level
C	Complex	Action objects not declared	Simple	Action declared
D	Simple	Action object not declared	Complex	Action not declared to enough quality level

performed either with an AR solution or with current alternatives such as emails or phone-calls. The messages from which tasks have to be inferred by technicians are presented in Table 5.

Table 6 presents the factor levels of these messages according to validation criteria (object (OA) and procedure (PA) awareness). Fig. 8 presents an example of the proposed solution to conduct the case study's operation.

The equipment used in the case study is a prototype of an aircraft's fuel hatch that was also manufactured by the company (Fig. 9). It includes components that are common in the aerospace industry such as rivets or panels. It has some pre-made defects which are common for the aerospace context such as dents (Fig. 9.d), scratches and breaks (Fig. 9.c). So, it can be said that the case study proposed is a fair representation of a common remote diagnosis operation in the aerospace industry. This prototype, resulted from the expert interviews, was later used in the laboratory experiments.

The sample size for the experiment can be calculated "a priori". In order to do it, an F-test for a three-way ANOVA experiment was applied. The test considers the number of groups to be 8 (factor levels), a variance of 0.2 (medium value of partial eta squared), a type-I error of 0.1 (alpha) and a power of 0.9 (1 – beta). The resulting required total sample size is 37 people.

A total of 30 MSc students (20 males and 10 females) participated as testers in laboratory experiments. Their ages range from 22 to 25 years and they are all enrolled on engineering-related degrees. Although, they have some basic knowledge in AR and maintenance due to their courses, they have no previous hands-on experience in any of them. The testers were given a short training on AR devices right before experimentation to avoid the presence of any learning curves. They were randomly allocated to one of the two experimental groups on AR usage factor (15 to AR and 15 to NOAR) to avoid carry-over effects on diagnosis procedures. Only those testers allocated to the AR group were given the post-experimental questionnaire.

A total of 8 real-life experts took part in the study for expert demonstrations. Their ages range from 20 to 60 years and they are all enrolled as part of the engineering repairs team within the company. Their roles were one of the following: structural design engineer, repairs engineer, and technical mechanical apprentice. Their experiences in those roles range from 2 to 22 years. None of them had any previous experience with AR. Before completing the usability and feasibility questionnaire, they were all given a short training on AR devices and then left to test both ends (expert and technician) of the proposed solution. The solution test occurred in an area within the company's facilities where the environmental conditions were considered similar to

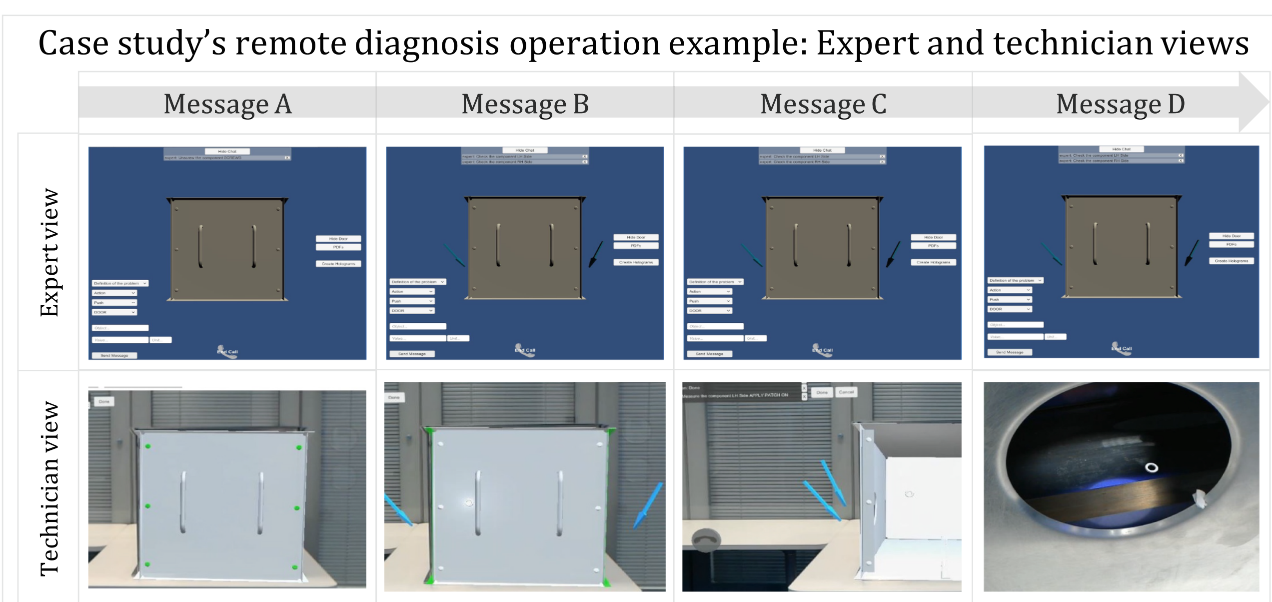

Fig. 8. Description of remote diagnosis case study AR-based example.

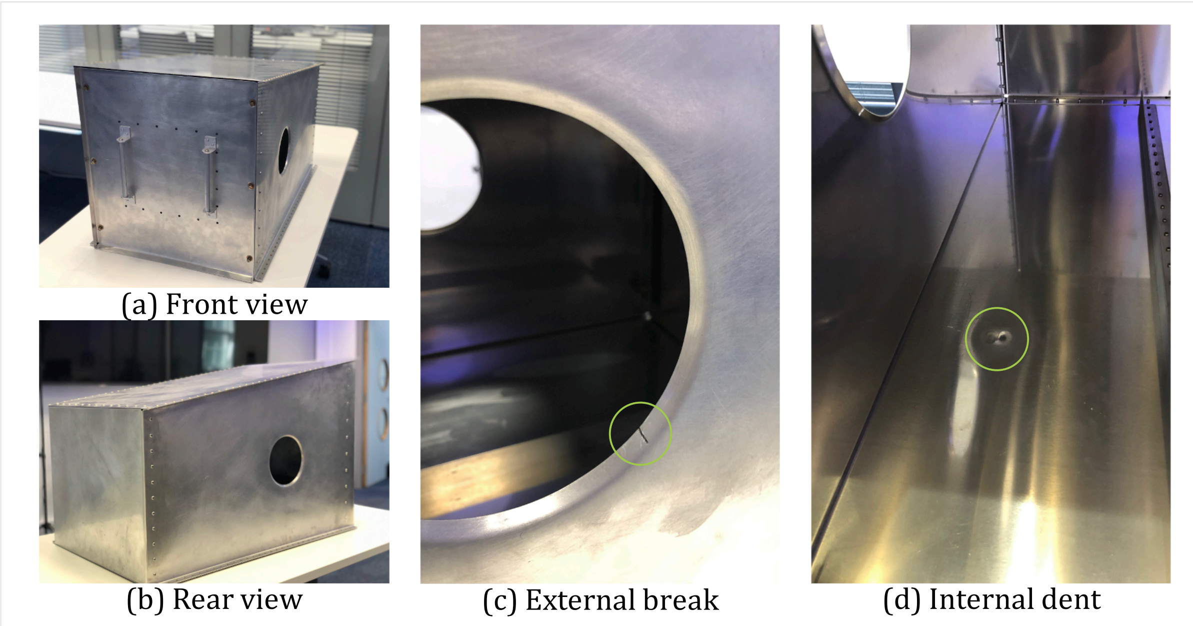

Fig. 9. Pictures of the aircraft's fuel hatch prototype and its defects.

those that occur in real-life scenarios (e.g. variable illuminance and confined spaces). So, it can be assumed that their judgements were based on real-life conditions from both, environmental and user perspectives.

The results of both, questionnaire as well as the stopwatch time and errors study, are presented in the following section.

6. Results and discussion

The aim of the experiments was to collect data to demonstrate the hypotheses presented in Section 5.2. The validity of these hypotheses is evaluated through the analysis of the experimental results presented in the following subsections.

6.1. Error results

Messages with different complexity levels were pre-defined for the experiments. Errors were defined as the number of tasks completed by a tester, which deviated in form or result of what is declared by the corresponding remote message. Hence, if the remote message were the same, then the number of errors should not be different between AR and NOAR solutions. As part of the analysis, it was necessary to test this assumption in order to ensure a direct correlation between experimental times and remote diagnosis efficiency.

Error results are presented in Fig. 10 and Table 7 shows the number of errors made by each tester in their experiments. The number of errors made by testers is low, with only 4 testers out of 30 committing more than one error. Moreover, if we consider the total number of tasks that the messages comprised (10), then the overall number is also very low as the maximum number of errors (3) made by a tester do not involve more than a third of the total tasks (33%). The results in Table 7 also indicate that there is no significant variance in the average number of results according to the solution used. The means and standard deviations on the number of errors for AR and NOAR solutions are the same for each group. Hence, the average number of errors made by a tester can be calculated at 6% (0.6/10).

6.2. Time results

Time results aim at evaluating the effect of utilising the proposed solution in remote diagnosis scenarios and its effect at different levels of message complexity. As time is considered to be a direct measure of the diagnosis efficiency, a reduction of time produced by the use of the solution proposed can be understood as an enhancement of remote diagnosis efficiency. Besides, it is also interesting to analyse the variation of such time reduction by the use of the research solution, if it exists, for different levels of remote message complexity. Message complexity is classified in terms of object and procedure awareness at four different levels:

1. **Simple-Simple (Message A (Table 5)):** message is simple in terms of object and procedure awareness.
2. **Simple-Complex (Message D):** message is simple in object awareness and complex in procedure awareness.
3. **Complex-Simple (Message C):** message is complex in object awareness and simple in procedure awareness.
4. **Complex-Complex (Message B):** message is complex in terms of object and procedure awareness.

Time results are presented in Fig. 11 and Tables 8–10. On average, the reduction on completion time of the AR solution compared to NOAR is 56% with average completion time per message using AR being 75.4 s and NOAR 134.0 s. Nevertheless, since other factors (object and procedure awareness) may also affect the time results, it seems valuable to evaluate the effect of each factor and calculate the difference in time between solutions for each factor level combination.

The box and whiskers plot presented in Fig. 11 visualises the variation in time for each solution classified by message complexity. It can be seen that for each complexity level, the AR solution has a shorter completion time than the NOAR solution. It can also be appreciated that the time gap increases with the complexity of the messages. Nonetheless, the variations shown in the graph cannot be compared directly since the tasks associated with each message cannot be assumed to have a similar completion time. Hence, the statistical significance of such variations requires further evidence to be considered valid. That is the reason why a three-way ANOVA analysis was conducted on the time variable according to the three main factors of the experiment: solution, object awareness, and procedure awareness. The ANOVA results are shown in Table 8.

The ANOVA results present evidence on the statistical significance

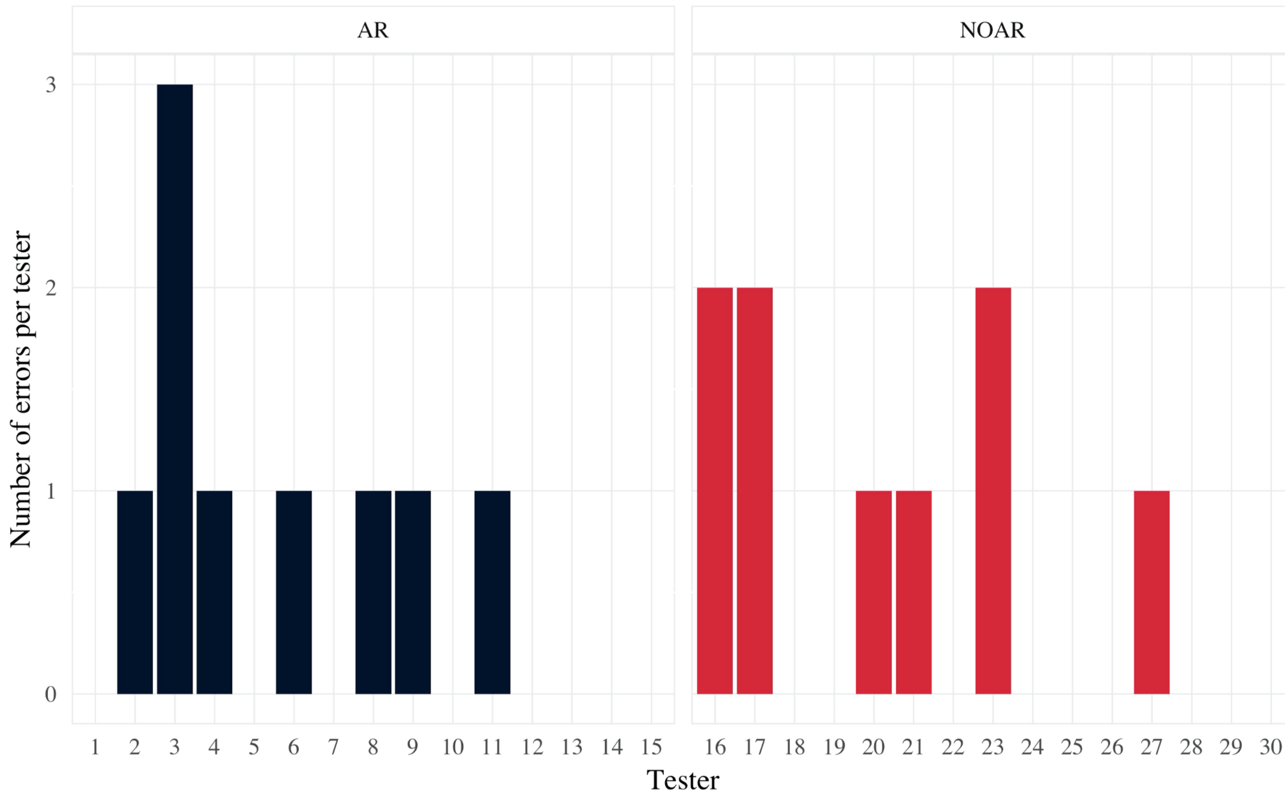

Fig. 10. Results on completion errors organized by experimental groups.

Table 7
Mean and standard deviations of experimental groups on number of completion errors.

Experiment group	Number of testers	Mean	Standard deviation
AR	15	0.6	0.828
NOAR	15	0.6	0.828

of the effect of the experimental factors and their interactions on the dependent variable (time). For a confidence interval of 95% (p-value equal or less than 0.05), it can be said that each factor, as well as their first level interactions, have a significant effect on completion time.

The effects and interactions between factors on the time variable can be further analysed by evaluating the difference in time reduction between solutions for each message complexity level (Table 9). For level "simple-simple" (message A), the AR solution shows a time reduction of 13% on average. For the "simple-complex" level (message D), time reduction for the AR compared to the NOAR solution is calculated at 32%. For "complex-simple" (message C) and "complex-complex" (message B) levels, time reductions are averaged at 54% and 59%, respectively.

Post hoc comparisons results from the Tukey HSD test can reveal significant differences between different factor groups. Table 10 shows the p-values for group factors interactions. Differences can be considered significant when p-values are equal or less than 0.05, with a confidence interval of 95%.

Significant differences can be found between AR and NOAR solutions for messages with complex object awareness, complex procedure awareness or a combination of both complexities. When procedure and object awareness are both simple, differences cannot be considered significant. Hence, it can be said that the higher the complexity of messages the better improvements AR solutions provide in terms of time reduction; and so, enhanced efficiency.

6.3. Usability and feasibility results

Questionnaires aimed at quantifying testers' and experts' opinions on the proposed solution usability, and experts' opinions on its feasibility for real-life conditions of remote diagnosis operations. Graphical and numerical results of these questionnaires can be seen in Fig. 12 and Table 11, respectively.

In terms of usability, experts and testers were questioned about the ease of use, ease to learn and overall satisfaction when using the proposed solution as technicians. According to results, both groups agree that the solution proposed can be easily used by technicians. They all scored above 4 (out of 5) for every criterion.

Additional usability results are related to the usability of each tool, provided by the solution proposed. Experts and testers were also questioned about the utility of the proposed AR solution. Although, both groups agree the solution proposed is usable, the scores are lower on average compared to the usability criterions. This correlates with some of the comments provided by experts and testers who mentioned the need to provide more "professional-looking" interfaces that can make the communication "smoother".

From a quality perspective, it was found interesting to ask testers and experts their expectations on how much the proposed solution could improve current remote diagnosis performances. According to the results presented in Fig. 12 and Table 11, they both coincide that time could be reduced on average around ~50%. That is a similar result to the time reduction evidenced by the experiments.

The importance of the feasibility questionnaire was to determine whether real-life experts might consider that the solution proposed could be used in real-life remote diagnosis scenarios. In order to do so, they were asked about the ability of the AR solution to be used in such scenarios and also in the ability of the data recorded to be re-used in other contexts (data usage). All experts agreed that the four tools (button interaction, message log, picture interaction and user interface) could be used by real-life technicians. Besides, only two of them

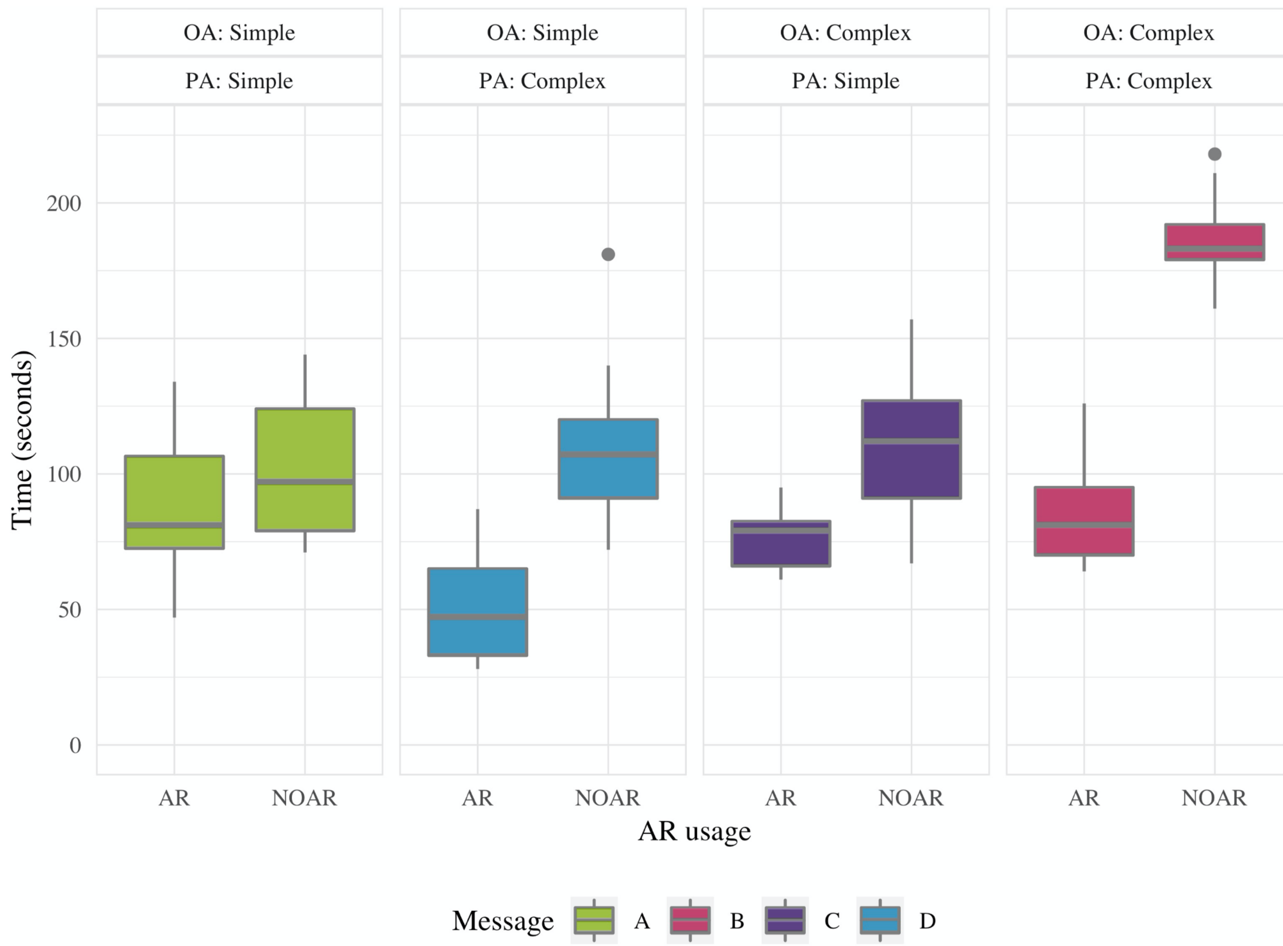

Fig. 11. Results on completion times organised by experimental groups and messages.

Table 8
Summary of results of three-way ANOVA test over completion time.

Effect	Df	Sum Sq	Mean Sq	F value	Pr (> F)	Significant (95% ci)
Solution	1	103,312	103,312	99.401	2.00e − 16	Yes
OA	1	34,782	34,782	33.465	6.69e − 18	Yes
PA	1	11,505	11,505	11.070	1.18e − 03	Yes
Solution:OA	1	15,436	15,436	14.852	1.94e − 04	Yes
Solution:PA	1	34,578	34,578	33.269	7.23e − 08	Yes
OA:PA	1	38,128	38,128	36.684	1.90e − 08	Yes
Solution:OA:PA	1	3956	3956	3.806	0.05356	No
Residuals	112	116,406	1039	—	—	—

Table 9
Mean and standard deviations of experimental groups on completion times.

Message	AR usage	Object A.	Procedure A.	No. of testers	Mean	Std. deviation
A	AR	Simple	Simple	15	88.7	24.2
	NOAR	Simple	Simple	15	102.2	25.7
D	AR	Simple	Complex	15	50.1	19.9
	NOAR	Simple	Complex	15	109.4	26.8
C	AR	Complex	Simple	15	75.9	11.2
	NOAR	Complex	Simple	15	112.5	25.8
B	AR	Complex	Complex	15	85.7	20.5
	NOAR	Complex	Complex	15	212.6	68.9

disagreed with the idea that the message-related data recorded (data usage) could be later reused in other areas (e.g. repair design) or in other remote diagnosis operations.

6.4. Discussion

The objective of the analysis results presented in the previous subsection is provide enough evidence to validate the hypotheses presented at the introduction of Section 6.

First, the error results offered sufficient evidence to say that within the sample there is no significant variation in the number of errors when using the AR solution proposed or an alternative. Hence, it can be considered that the assumption about errors being invariable on the effect of the solution is valid. And so, completion time can be used as a direct measure of diagnosis efficiency. Nevertheless, the number of errors was counted for the whole experiment and not per message. Therefore, it could not be studied whether the level of complexity of messages had an effect on the appearance of errors. Although this might be an interesting element to evaluate, it was out of this study's scope as the messages were pre-determined in order to analyse the effect of the proposed solution at different levels of message complexity. Future studies could research into this aspect of the effect of AR solutions in the context of remote diagnosis.

Second, the time results presented enough and significant evidence to evaluate the correlation between the effect of the AR solution and the complexity of messages in the efficiency of remote diagnosis. The ANOVA results and consequent post hoc comparisons showed a significant effect of the proposed solution over the completion time and a significant correlation with the effect of message complexity in remote diagnosis efficiency. These results emphasised the ability of the proposed solution to enhance situational awareness of on-site technicians increased with the complexity of messages being sent by remote experts. On average, the effect on the proposed solution on time reduction

Table 10
Significance (p-value) results from post hoc comparisons Turkey HSD test.

	AR-SS	AR-SC	AR-CS	AR-CC	NOAR-SS	NOAR-SC	NOAR-CS	NOAR-CC
AR-SS	—	0.0296	0.9582	0.9999	0.944	0.6915	0.5098	0
AR-SC	0.0296	—	0.3681	0.0605	0.0006	0.0001	0.0001	0
AR-CS	0.9582	0.3681	—	0.9908	0.3382	0.1104	0.0552	0
AR-CC	0.9999	0.0605	0.9908	—	0.8534	0.5213	0.348	0
NOAR-SS	0.944	0.0006	0.3382	0.8534	—	0.9994	0.9919	0
NOAR-SC	0.6915	0.0001	0.1104	0.5213	0.9994	—	0.9999	0
NOAR-CS	0.5098	0.0001	0.0552	0.348	0.9919	0.9999	—	0
NOAR-CC	0	0	0	0	0	0	0	—

was of 56% compared to current alternative solutions (phone calls and emails). Besides, this reduction of completion time was higher for those messages with higher levels of complexity (from 13% on the simplest one to 59% on the most complex one). Hence, it can be said that the effect of the proposed solution on remote diagnosis efficiency is worthier in the cases when the complexity of messages is higher. Nonetheless, the implementation of such solution may not be worthy in terms of efficiency gains due to the smaller effect found in the

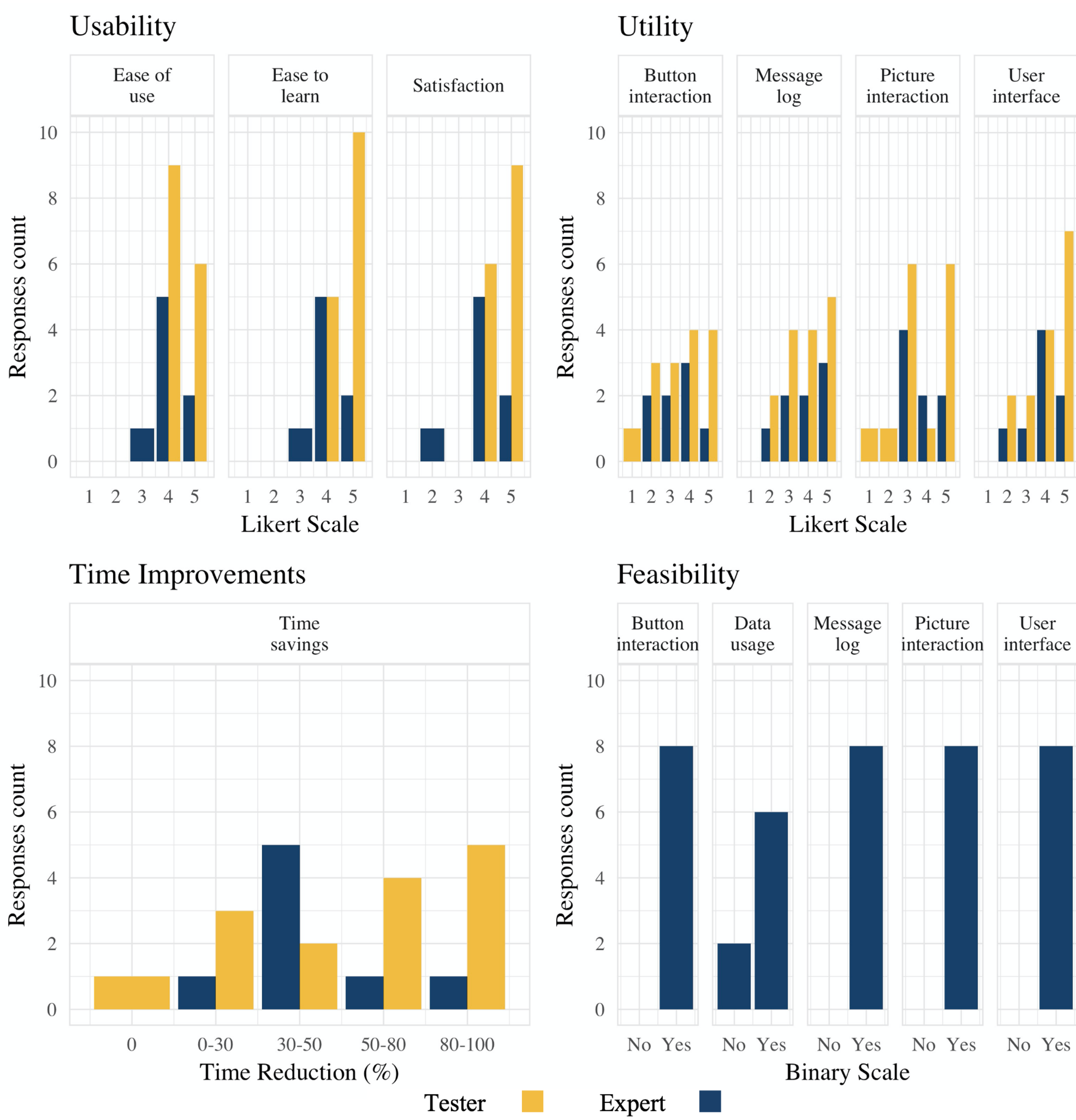

Fig. 12. Results on usability and feasibility questionnaires for experimental testers and real-life experts.

Table 11
Means and standard deviations on questionnaire results by interviewees.

Criterion	Element	Interviewee	Responses	Mean	Std. deviation
Usability	Ease of use	Tester	15	4.40	0.507
		Expert	8	4.12	0.641
	Ease to learn	Tester	15	4.67	0.488
		Expert	8	4.12	0.641
	Satisfaction	Tester	15	4.60	0.507
		Expert	8	4.00	0.926
Utility	Button interaction	Tester	15	3.47	1.300
		Expert	8	3.38	1.060
	Picture interaction	Tester	15	3.67	1.290
		Expert	8	3.75	0.886
	User interface	Tester	15	4.07	1.100
		Expert	8	3.88	0.991
	Message log	Tester	15	3.80	1.080
		Expert	8	3.88	1.130
Quality	Time reduction	Tester	15	55.70	32.500
		Expert	8	46.20	22.200
Feasibility	Button interaction	Expert	8	1.00	0.000
	Picture interaction	Expert	8	1.00	0.000
	User interface	Expert	8	1.00	0.000
	Message log	Expert	8	1.00	0.000
	Data usage	Expert	8	0.75	0.463

experiments. Besides, the experiments were conducted on the assumption of pre-determined messages. The ability of the proposed solution to simplify the labour of remote experts when sending messages was not studied in detail. Future studies could investigate the effect of such solutions on the remote expert side. They could also look into the relation between efficiency gains and the actual decrease of related costs to analyse the worthiness of implementing such solutions in real-life scenarios.

Finally, the usability and feasibility questionnaires provided enough evidence to say that in the opinion of testers and real-life experts, the solution proposed is useful and it could be used in real-life scenarios. The results given by testers and experts were similar, showing a consistent trend that could further validate the conclusions extracted from the results. An important aspect included in the questionnaire results was the evaluation of the solution's ability to collect diagnosis-related data that could be further used in other maintenance-related areas. Although experts' opinion validated such idea, future research should focus on demonstrating the ability of expert-data captured to be re-used in other areas such as maintenance planning, training or recommendations for future calls. Such investigations should analyse the nature of the data being collected and propose and validate different methods that could make use of it to enhance the abovementioned process (e.g. performance or prognostics).

An important aspect to note regarding results validity involves the experimental sample size. The "a priori" F-test conducted to determine the sample size required for the three-way ANOVA determined a total of 37 people. Although the final number of testers was 30, the final sample size is close enough to the "a-priori" calculated sample size to still consider the results valid. Besides, similar researches reviewed [6,46,48] also utilised a similar number of testers. Hence, the results of the questionnaires can also be considered significant according to the sample size.

Another impact aspect to note regarding results validity relates to the experimental design. Experiments purpose was to analyse on-site technician's efficiency improvements in remote diagnosis when using AR structured authoring methods to enable remote collaboration. So, it seems necessary to include as relevant analysis factors those affecting remote collaboration. Literature [15-19] suggests Object (OA) and Procedure Awareness (PA) as the most relevant, which further depend on the complexity of the message being sent. Hence, the validation experiments were specifically designed to measure the effects of AR structured authoring in diagnosis efficiency according to OA, PA and message complexity. The presented statistical analysis tests all possible factor combinations, showing the effects of structured communication in OA and PA in simple and complex cases. Its results suggest that structured communication is more efficient compared to un-structured, with higher efficiency increases in cases of more complex messages in terms of OA and PA (Table 9). In order to collect data for analysing the abovementioned effects, it was necessary to choose an alternative unstructured approach for remote communication. The decision to choose a combination of phone calls and emails as the alternative was based on the two reasons. First, it met the requirements of being an un-structured approach. Second, this alternative is a common approach used by real-life experts in their daily jobs. So, the surveyed experts regarding usability and feasibility of the proposed AR method could provide more accurate responses. Nevertheless, other alternative approaches could be subject of comparison if efficiency improvements on the expert side were to be analysed. Un-structured AR methods [26,33] or alternative visualisations such as 3D-based PDF documents [49] are also relevant for comparison because of different reasons.

In the case of un-structured AR authoring methods, experts are free to consider the augmentation method (AR content type) that better suits a specific message. This will involve analysing other effects rather than OA and PA since arbitrary actions can be taken from expert testers, which are strongly dependant on testers capabilities. The authors estimate that such validation would require a bigger sample size from both, experts' and technicians' testers, of an order of magnitude higher at the AR proposal's current implementation.

In the case of alternative visualisation methods, such as 3D-based PDF documents, these involve additional content types (e.g. 3D models on a screen). Evaluation of alternative content types would require taking into consideration different effects on remote collaboration such as ergonomic factors. For example, screen location in the on-site technician's view while conducting remote collaborative diagnosis. The authors estimate that including these factors within the analysis may require again a bigger sample size. Besides, this sample size would also be increased if additional effects of un-structured communication were to be considered.

Another relevant aspect regarding alternative approaches to compare in experiments is the use of video streaming. In the proposed solution, video streaming the technician's view to the remote expert's view was an option to confirm correct remote message execution. Although separated from the expert's 3D environment (Fig. 7), it was used during the experiments. The use of video streaming introduces other alternatives for experimental comparison. These alternatives are those of remote video guidance and can be of two types: with and without AR support. For the first type, comparisons should be made against an AR un-structured authoring approach to comply with the experiments aim, which has been discussed above. For the second type, such alternative would be similar to the case of phone and emails because messages could be sent solely through text or audio. This alternative could have positive (e.g. faster message confirmation) and negative effects (e.g. slower message execution due to video streaming issues) on diagnosis efficiency compared to phone and emails. The authors considered that these negative effects could have a higher impact than the positive ones. Hence, this alternative comparison may positively discriminate the proposed solution. Besides, video streaming can have further impacts on real-life working environments (e.g. latency due to extended bandwidth usage). Hence, further experiments could be made in real-life conditions to analyse such effects.

In the discussion above, there are a number of factors (e.g. ergonomics, video streaming bandwidth usage and AR content types suitability) that are very relevant for analysis. And the current AR structured authoring approach would benefit if taken into consideration. For example, the current proposal could be further improved if recommendations based on diagnosis context were given to experts regarding following messages to be sent to on-site technicians. Such recommendations could be then further tested against un-structured AR approaches with no recommendation capabilities. Also, the current proposal could also be enhanced by providing dynamic adaptation of content types to the most suitable ones for technicians regarding their expertise and other ergonomic factors. For example, using overlaid 3D models when the message sent requires operation of a specific component, while providing non overlaid, 3D exploded equipment views when messages require holistic understanding of it. Such dynamic adaptability could be then further tested against other alternative visualisation methods with fixed authoring methods. Besides, alternative means for message execution confirmation other than video streaming could be used to reduce bandwidth usage of the proposed solution. These suggested improvements would require of experimentation in real-life working environments to prove their benefits and feasibility. Hence, the authors consider them very relevant for later stages of this research.

7. Conclusions and future works

7.1. Conclusions

The paper (1) defines a structure that regulates message elements to ensure their situational awareness, and (2) develops a rule-based authoring approach that automatically creates AR content of message elements to enhance remote diagnosis efficiency. The proposed AR-based communication contributes to filling an important research gap in the remote diagnosis context regarding improvements on object and procedure awareness of AR-based remote collaboration. The developed system offers structured and real time communication methods using automated augmentation of message elements. The validation results indicate strongly that AR based technology can improve the efficiency in remote communication in terms of time and errors reduction, whilst also being efficient and valid in working environments.

7.2. Future works

Future works will explore how remote communication can be enhanced beyond the diagnosis focus in this paper. A relevant area to consider is the scope of communication to includes those cases when it is not simply bi-directional (e.g. one-to-many or many-to-one). For that, it seems necessary to further analyse the aspects that involve effective reception of a message and the consequences on the communication protocol. It also includes the need to analyse alternative AR content types that could be implemented (e.g. voice-to-text). Besides message reception, another relevant aspect is message execution confirmation. Although video streaming is used in the proposed solution to achieve that, alternative AR content types can also be useful for that purpose. They can provide additional benefits such as reduced bandwidth usage and so, faster communication. These future works, along with those suggested in forthcoming paragraphs, would require of further experiments in real-life working environments. This is required to consider all possible factors (e.g. ergonomics, expertise, network range, etc.) that can have an effect on diagnosis efficiency.

Another relevant area is that of dynamic content adaption regarding ergonomic and expertise level factors. Such work would involve analysis of relevant factors on remote collaboration and development of additional methods to capture necessary contextual data and dynamically adapting AR content types to it. Besides, it would also involve further experimentation to compare the improved AR structured authoring proposal against alternative visualisation methods such as 3D-based PDF documents [49].

Finally, another important aspect to consider is the potential long-term improvements that structured communication can provide. Structured communication can simplify the process of further analysing historical messages. It can be used to develop recommender systems that make use of historical messages and contextual data to further help remote experts when collaborating with on-site technicians. Further experimentation would also be necessary to compare such developments against other un-structured AR approaches that still make use of AI methods to help on-site technicians in early diagnosis stages [36].

Declaration of Competing Interest

None.

Acknowledgements

This research was partially supported by the EPSRC funded project DigiTOP, grant number EP/R032718/1, and the Slovenian Research Agency, grant no. P2-0270. It was a cooperative work between Cranfield (UK) and Ljubljana (Slovenia) Universities.

References

[1] S.C.Y. Lu, W. Elmaraghy, G. Schuh, R. Wilhelm, A scientific foundation of collaborative engineering, CIRP Ann. – Manuf. Technol. (2007), https://doi.org/10.1016/j.cirp.2007.10.010.
[2] S.K. Ong, Y. Shen, A mixed reality environment for collaborative product design and development, CIRP Ann. – Manuf. Technol. (2009), https://doi.org/10.1016/j.cirp.2009.03.020.
[3] G. Dini, M. Dalle Mura, Application of Augmented Reality Techniques in Through-life Engineering Services. Procedia CIRP, vol. 38, Elsevier B.V.; 2015, p. 14–23. doi:10.1016/j.procir.2015.07.044.
[4] Global Remote Asset Management Market Analysis & Trends - Industry Forecast to 2025, 2017.
[5] R. Roy, R. Stark, K. Tracht, S. Takata, M. Mori, Continuous maintenance and the future - Foundations and technological challenges, CIRP Ann. – Manuf. Technol. 65 (2016) 667–688, https://doi.org/10.1016/j.cirp.2016.06.006.
[6] X. Wang, S.K. Ong, A.Y.C. Nee, Multi-modal augmented-reality assembly guidance based on bare-hand interface, Adv. Eng. Inform. 30 (2016) 406–421, https://doi.org/10.1016/j.aei.2016.05.004.
[7] R. Masoni, F. Ferrise, M. Bordegoni, M. Gattullo, A.E. Uva, M. Fiorentino, et al., Supporting remote maintenance in Industry 4.0 through augmented reality, Proc. Manuf. (2017), https://doi.org/10.1016/j.promfg.2017.07.257.
[8] A.Y.C. Nee, S.K. Ong, G. Chryssolouris, D. Mourtzis, Augmented reality applications in design and manufacturing, CIRP Ann. – Manuf. Technol. 61 (2012) 657–679, https://doi.org/10.1016/j.cirp.2012.05.010.
[9] J.A. Erkoyuncu, I. Fernández del Amo, M. Dalle Mura, R. Roy, G. Dini, Improving efficiency of industrial maintenance with context aware adaptive authoring in augmented reality, CIRP Ann. – Manuf. Technol. 66 (2017) 465–468, https://doi.org/10.1016/j.cirp.2017.04.006.
[10] I. Fernández Del Amo, J.A. Erkoyuncu, R. Roy, R. Palmarini, D. Onoufriou, A systematic review of Augmented Reality content-related techniques for knowledge transfer in maintenance applications, Comput. Ind. 103 (2018) 47–71, https://doi.org/10.1016/j.compind.2018.08.007.
[11] A. Capodieci, L. Mainetti, L. Alem, An innovative approach to digital engineering services delivery: An application in maintenance, in: Proc. - 2015 11th Int. Conf. Innov. Inf. Technol. IIT 2015, Institute of Electrical and Electronics Engineers Inc.; 2016, p. 342–9. doi:10.1109/INNOVATIONS.2015.7381565.
[12] L. Alem, Supporting the changing roles of maintenance operators in mining: a human factors perspective, Ergon. Open J. 4 (2011) 81–92, https://doi.org/10.2174/1875934301104010081.
[13] E. Poppe, R. Brown, J. Recker, D. Johnson, I. Vanderfeesten, Design and evaluation of virtual environments mechanisms to support remote collaboration on complex process diagrams, Inf. Syst. 66 (2017) 59–81, https://doi.org/10.1016/j.is.2017.01.004.
[14] M. Germani, M. Mengoni, M. Peruzzini, An approach to assessing virtual environments for synchronous and remote collaborative design, Adv. Eng. Inform. 26 (2012) 793–813, https://doi.org/10.1016/j.aei.2012.06.003.
[15] J. Oyekan, V. Prabhu, A. Tiwari, V. Baskaran, M. Burgess, R. Mcnally, Remote real-time collaboration through synchronous exchange of digitised human–workpiece interactions, Futur. Gener. Comput. Syst. 67 (2017) 83–93, https://doi.org/10.1016/j.future.2016.08.012.
[16] S.R. Fussell, L.D. Setlock, J. Yang, J. Ou, E. Mauer, Gestures over video streams to support remote collaboration on physical tasks, Human-Comput. Interact. 19 (2004) 273–309, https://doi.org/10.1207/s15327051hci1903_3.
[17] N.V. Flor, Side-by-side collaboration : a case study, Int. J. Hum. Comput. Stud. 49 (1998) 201–222, https://doi.org/10.1006/ijhc.1998.0203.
[18] C.E. Ford, Collaborative construction of task activity: coordinating multiple resources in a high school physics lab, Res. Lang. Soc. Interact. 32 (1999) 369–408, https://doi.org/10.1207/S15327973rls3204_3.
[19] C.T. John, Findings from observational studies of collaborative work, Int. J. Man Mach. Stud. 34 (1991) 143–160, https://doi.org/10.1016/0020-7373(91)90039-A.
[20] R.E. Kraut, S.R. Fussell, J. Siegel, Visual Information as a Conversational Resource in Collaborative Physical Tasks. vol. 18, 2003.
[21] O. Fakourfar, K. Ta, R. Tang, S. Bateman, A. Tang, Stabilized Annotations for Mobile

Remote Assistance, Association for Computing Machinery (ACM), 2016, pp. 1548–60. doi:10.1145/2858036.2858171.
[22] D. Gelb, A. Subramanian, K.-H. Tan, Augmented reality for immersive remote collaboration. IEEE Work. Pers. Vis., Kona, HI, USA: 2011. doi:10.1109/POV.2011.5712368.
[23] R.T. Azuma, The Most Important Challenge Facing Augmented Reality. Presence Teleoperators Virtual Environ., 25 (2016) 234–238. doi:10.1162/PRES_a_00264.
[24] E. Bottani, G. Vignali, Augmented reality technology in the manufacturing industry: a review of the last decade, IISE Trans. 51 (2019) 284–310, https://doi.org/10.1080/24725854.2018.1493244.
[25] G. Dini, Mura M. Dalle, Application of augmented reality techniques in through-life engineering services, Proc. CIRP 38 (2015) 14–23, https://doi.org/10.1016/j.procir.2015.07.044.
[26] P. Gurevich, J. Lanir, B. Cohen, Design and Implementation of TeleAdvisor: a projection-based augmented reality system for remote collaboration, Comput. Supp. Coop Work CSCW An Int. J. 24 (2015) 527–562, https://doi.org/10.1007/s10606-015-9232-7.
[27] S. Fox, The importance of information and communication design for manual skills instruction with augmented reality, J. Manuf. Technol. Manag. 21 (2010) 188–205, https://doi.org/10.1108/17410381011014369.
[28] M. Adcock, C. Gunn, Using projected light for mobile remote guidance, Comput. Supp. Coop. Work CSCW An Int. J. 24 (2015) 591–611, https://doi.org/10.1007/s10606-015-9237-2.
[29] O. Oda, C. Elvezio, M. Sukan, S. Feiner, B. Tversky, Virtual replicas for remote assistance in virtual and augmented reality, Assoc. Comput. Mach. (ACM) (2015) 405–415, https://doi.org/10.1145/2807442.2807497.
[30] N. Zenati-Henda, A. Bellarbi, S. Benbelkacem, M. Belhocine, Augmented reality system based on hand gestures for remote maintenance, Int. Conf. Multimed. Comput. Syst., Marrakech, Morroco (2014), https://doi.org/10.1109/ICMCS.2014.6911258.
[31] J. Rambach, A. Pagani, M. Schneider, O. Artemenko, D. Stricker, 6DoF Object Tracking based on 3D Scans for Augmented Reality Remote Live Support, Computers 7 (2018) 6, https://doi.org/10.3390/computers7010006.
[32] K.P.K. Reddy, B. Venkitesh, A. Varghese, N. Narendra, G. Chandra, P. Balamuralidhar, Deformable 3D CAD models in mobile augmented reality for tele-assistance, in: Proc. - APMediaCast 2015 Asia Pacific Conf. Multimed. Broadcast., Institute of Electrical and Electronics Engineers Inc., 2015, pp. 51–5. doi:10.1109/APMediaCast.2015.7210274.
[33] D. Mourtzis, A. Vlachou, V. Zogopoulos, Cloud-based augmented reality remote maintenance through shop-floor monitoring: a product-service system approach, J. Manuf. Sci. Eng. 139 (2017) 061011, , https://doi.org/10.1115/1.4035721.
[34] K. El Ammari, A. Hammad, Remote interactive collaboration in facilities management using BIM-based mixed reality, Autom. Constr. 107 (2019) 102940, , https://doi.org/10.1016/j.autcon.2019.102940.
[35] M. Neges, C. Koch, M. König, M. Abramovici, Combining visual natural markers and IMU for improved AR based indoor navigation, Adv. Eng. Inform. 31 (2017) 18–31, https://doi.org/10.1016/j.aei.2015.10.005.
[36] E. Hadar, J. Shtok, B. Cohen, Y. Tzur, L. Karlinsky, Hybrid remote expert - an emerging pattern of industrial remote support, in: CEUR Workshop Proc., vol. 1848, CEUR-WS, 2017, pp. 33–40.
[37] A. Deshpande, I. Kim, The effects of augmented reality on improving spatial problem solving for object assembly, Adv. Eng. Inform. 38 (2018) 760–775, https://doi.org/10.1016/j.aei.2018.10.004.
[38] I.F. Del Amo, E. Galeotti, R. Palmarini, G. Dini, J. Erkoyuncu, R. Roy, An innovative user-centred support tool for Augmented Reality maintenance systems design: a preliminary study, Proc. CIRP 70 (2018) 362–367, https://doi.org/10.1016/j.procir.2018.02.020.
[39] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, J. Manag. Inf. Syst. 24 (2008) 45–77, https://doi.org/10.2753/mis0742-1222240302.
[40] A. Booth, D. Papaioannou, A. Sutton, Systematic Approaches to a Successful Literature Review, Sage Publications, Thousand Oaks, CA, USA, 2012.
[41] J.P. Zachman, The Zachman framework evolution. Zachman Int. Inc., 2011, doi:10.1038/469299b.
[42] J.A. Erkoyuncu, I. Fernández del Amo, M. Dalle Mura, R. Roy, G. Dini, Improving efficiency of industrial maintenance with context aware adaptive authoring in augmented reality, CIRP Ann. – Manuf. Technol. (2017), https://doi.org/10.1016/j.cirp.2017.04.006.
[43] C. Robson, K. McCartan, Real World Research, fourth ed., Wiley, Hoboken, NJ, USA, 2016.
[44] J. Nielsen, Usability Engineering, first ed., Morgan Kaufmann, San Francisco, California, USA, 1993.
[45] H.L. Chi, Y.C. Chen, S.C. Kang, S.H. Hsieh, Development of user interface for tele-operated cranes, Adv. Eng. Inform. 26 (2012) 641–652, https://doi.org/10.1016/j.aei.2012.05.001.
[46] J. Gimeno, P. Morillo, J.M. Orduña, M. Fernández, A new AR authoring tool using depth maps for industrial procedures, Comput. Ind. 64 (2013) 1263–1271, https://doi.org/10.1016/j.compind.2013.06.012.
[47] R. Palmarini, J.A. Erkoyuncu, R. Roy, An Innovative Process to Select Augmented Reality (AR) Technology for Maintenance, Proc. CIRP, vol. 59, Cranfield, 2017, pp. 23–28. doi:10.1016/j.procir.2016.10.001.
[48] J. Zhu, S.K. Ong, A.Y.C. Nee, An authorable context-aware augmented reality system to assist the maintenance technicians, Int. J. Adv. Manuf. Technol. 66 (2013) 1699–1714, https://doi.org/10.1007/s00170-012-4451-2.
[49] F. Klier, A. Pfouga, S. Rulhoff, J. Stjepandić, An approach for agile planning of disassembly and recycling of complex products, Int. J. Agil. Syst. Manag. 10 (2017) 250–270, https://doi.org/10.1504/ijasm.2017.088535.