Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 180 (2021) 269–279

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Industry 4.0 and Smart Manufacturing

Creating an Open-Source Augmented Reality Remote Support Tool for Industry: Challenges and Learnings

Andrea Aschauer*, Irene Reisner-Kollmann, Josef Wolfartsberger

University of Applied Sciences Upper Austria, Wehrgrabengasse 1-3, 4400 Steyr, Austria

Abstract

Remote support tools offer great savings potential for maintenance work and simplify communication between on-site workers and remote experts in the back-office. Despite the predicted benefits and the existence of an increasing number of commercial augmented reality (AR) remote support tools, the technology still struggles with various challenges, mainly caused by technical limitations, safety concerns or industry-specific constraints. In this paper, we describe the implementation of an open source AR remote support application with a strong industry focus. The software was developed based on the feedback of four industry partners with different needs in the field of maintenance work. Therefore, we set the focus on configurability to provide the right tools for each use case. The developed remote support tool was then compared with traditional support measures in a user study. This paper presents the lessons learned from the expert discussions with our company partners and the results of the user study.

Keywords: remote support; augmented reality; smart maintenance

1. Introduction

Maintenance management makes up an essential part of the annual operational budget. Therefore, the savings potential through increased efficiency is correspondingly large. Many companies are using maintenance tools to cut down running costs in this field. The three main approaches to conduct and document maintenance tasks are paper reports (39%), spreadsheets (52%) and computerized maintenance management systems (CMMS) (63%) [13]. The increased performance of mobile devices opens up new use cases for AR-supported remote maintenance, where a remote expert annotates the on-site worker's view with virtually anchored information. An increasing number of commercial AR remote maintenance applications like Microsoft Remote Assist or Vuforia Chalk were published in recent years. Even if the companies are interested in remote maintenance apps, reports show that only 9% of companies currently use mobile devices as part of the CMMS maintenance process [3]. Together with company partners in a

* Corresponding author.
E-mail address: Andrea.Aschauer@fh-ooe.at (Andrea Aschauer).

10.1016/j.procs.2021.01.164

collective research project, we analyzed the reasons for the lack of use of AR remote maintenance apps in industry. Commercial AR remote support tools often cannot be integrated into existing CMMS processes and trigger security concerns at management level. For our corporate partners, the lack of adaptability to existing business processes, such as integration with the current support ticket system, and strict internal security policies are the main reasons why existing AR-based remote support solutions are not an option for them. Based on our partner companies' feedback, we developed a prototype, which was compared with traditional approaches in a controlled user study.

The paper includes the following main contributions: (1) The implementation of an open source AR remote support application out of free software components. (2) Best-practice solutions for common problems. (3) A discussion about the use case specific differences based on a preliminary user study.

2. Related Work

Our implementation is based on a wide range of existing remote support techniques. We focused on methods which can be easily incorporated by using standard software and hardware components. The following section sums up important research work in the respective fields.

Tele-pointers are a simple technique for remote assistance. A video of the on-site worker's environment is sent to the remote expert. The remote expert can directly draw on top of this video stream [5]. Usually the annotations automatically disappear again, because they become invalid as soon as the viewpoint of the video changes. The on-site worker sees the tele-pointer in real-time. As the annotation only depicts single points without further information, accompanying communication via the voice channel is crucial. This approach is especially useful for quick instructions to direct the on-site worker's attention to a specific object [6].

The validity of annotations is largely extended, if they are stabilized in the reconstructed world space. Point-based annotations such as spheres or billboards are positioned in the 3D world space and visualized with augmented reality [9]. Provided that the AR tracking is good enough, the annotations stay on the same position independent from the viewpoint.

More sophisticated instructions are possible with 2D drawings instead of single points. It's possible to freeze the video stream [12] and make freehand drawings or place text or 2D icons on it [10]. Some systems [11, 7] are limited to 2D spaces such as table tops. Then the 2D annotations can be unambiguously projected into world coordinates.

If arbitrary 3D spaces are supported, it is important to find a useful projection of 2D annotations into the real-world environment [8]. Plane-based projections use either a plane parallel to the camera plane at the time of creation or a reconstructed world plane. Another technique is spray paint where each sample point of the annotation is independently projected onto the reconstructed scene. In any case, there are problems if the viewpoint changes too much and the depth of a sample point does not conform with the object of interest. A different approach is to classify special drawings such as circles or arrows and interpret them as 3D hints based on the reconstructed 3D geometry [15].

Instead of video signals, it's possible to share information via single screenshots [14, 1]. In the SEMarbeta-system[2], the screen either shows the live video or the currently frozen screenshot for both users. The remote expert and optionally also the on-site worker can draw annotations on top of the screenshot. Multiple screenshots can be organized in a gallery [18] and are useful as documentation for future maintenance cases.

In our work, we evaluated these approaches and selected the most promising techniques for our solution. We especially focused on industrial environments and a feasible realization with standard software and hardware tools.

3. Application description

Our AR remote support solution connects a portable device of an on-site worker with an remote expert using a desktop application. The on-site worker starts the call by sharing a unique connection key with the remote expert, e.g. by e-mail. The remote expert uses this key to establish the connection with the on-site worker. Remote experts and on-site workers see a live video recorded by the device of the on-site worker and can extend this live video with AR annotations in 3D space to facilitate communication with visual labels (see Figure 1 (a) and (c)). In addition, the annotations can be displayed together with the original screenshot in a gallery (see Figure 1 (b)). Both users can switch between the gallery and the live video view at any time by clicking on the gallery icon in the upper screen bar.

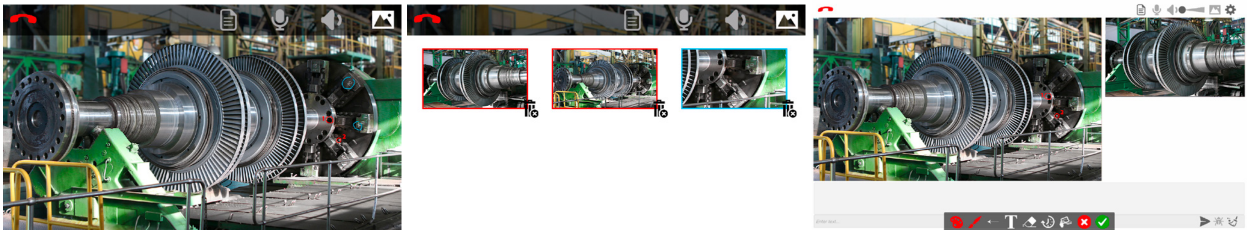

Fig. 1: Remote support application: (a) The application of the on-site worker shows the live camera video which is augmented by annotations in the AR 3D space. (b) The gallery overview on the device of t on-site device shows the snapshot at the time the annotation was created and a frame in the creator's default color. (c) The remote expert application shows the frozen live video that can be annotated with the tools from the opened drawing palette. The live video is displayed in a reduced form next to the drawing area.

Touching the video freezes the live image and opens the drawing tool. By closing this tool, the application switches back to the live video view, which immediately shows the new annotation in 3D space.

For the remote expert, we provide two tools for annotating the video: With a left click, the remote expert can make the mouse cursor visible in the live video of the on-site worker. A right click takes a screenshot of the current video frame and the drawing palette opens up. This second form of annotation is also available to the on-site workers. In addition to the still image, the live video is shown in a small window for the remote expert (see Figure 1 (c)). In this mode, the remote expert can use a freehand pen, an arrow tool, textual annotations and an eraser to modify drawings. Furthermore, the remote expert can change the drawing color. The default drawing colors are different for the remote expert (red) and the on-site worker (cyan). The gallery entries are also marked with a frame in these default colors to identify the first author of the annotation. The gallery entries can be viewed by both users, but only modified by the remote expert.

In addition to the live video, voice and message chat are available for communication with the remote expert. The chat option serves as a fallback solution for loud industrial environments where voice transmission is difficult. The mobile application for the on-site worker focuses on core feature in order to be easily operated. For the remote expert, a slightly more extensive range of functions is offered, such as session management by configuring different modes as described in Section 5.

4. System architecture

Our open source AR remote support application is based on free software packages. We used Unity 3D[1] as the application development environment, ARFoundation[2] for the calculation of the 3D space and the augmented reality view and WebRTC[3] for establishing communication between devices.

As shown in Figure 2, the remote support system is divided into three areas: remote expert view, data transmission and on-site worker view. The on-site worker is the initiator of the call and manages the real space. The remote expert view receives the information about the real space and augments this space with additional information. The data transmission layer forms the technical basis for the communication.

We use a WebRTC-plugin[4] to establish the real-time connection. It offers four communication channels: live video, audio, text and data. In our application, voice, chat and status messages are sent bidirectionally. The live video, still images and gallery entries are only broadcast from on-site workers to remote experts, while configuration settings and annotations are only sent in the opposite direction.

There are three main application layers on the on-site worker application which setup the augmented reality view. The basis is the ARFoundation-framework, which calculates the camera pose, 3D points and dominant planes. The

[1] https://unity.com/
[2] https://unity.com/unity/features/arfoundation
[3] https://webrtc.org/
[4] https://www.because-why-not.com/webrtc

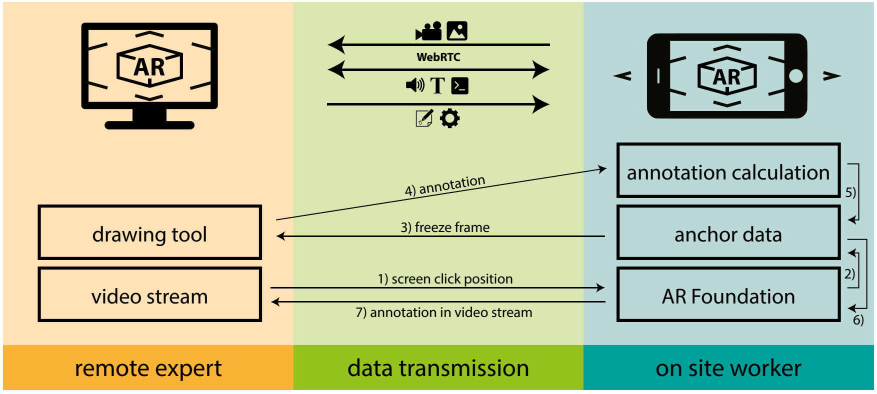

Fig. 2: System architecture and cooperation between the areas of remote experts, on-site workers and data transmission.

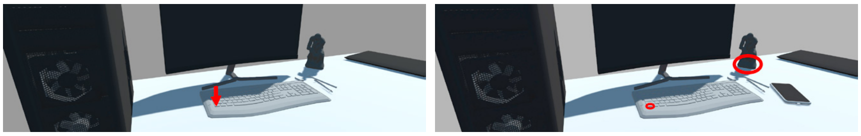

Fig. 3: Visualisation of the mouse cursor (left) vs. a drawing overlay (right).

anchor point layer manages the positioning of all virtual objects that expand the real space. The third layer is the annotation calculation. This layer transforms the 2D annotations into the 3D AR space.

All of the layers mentioned above are required when creating an annotation. The remote expert watches the live video. (1) By clicking on the video, the relative click position within the video is transmitted to the on-site worker. The on-site worker device takes this position and sends a raycast via ARFoundation to the reconstructed 3D scene of the real environment. (2) For this 3D position a new anchor is created and the management data for the anchor is stored in the anchor point layer. (3) The anchor point layer reports the successful creation of the anchor and its ID back to the remote expert. The remote expert video freezes and the drawing pallet appears. (4) The remote expert annotates the still frame and sends the drawing to the on-site worker device. Using the annotation calculation module, the on-site worker device transfers the 2D annotation into the 3D space. (5) The converted data are connected with the anchor ID and stored in the anchor point layer. (6) Finally this converted annotation is positioned in the reconstructed AR 3D space and (7) is visualized via ARFoundation in the real video stream, which is displayed on the on-site worker's device and sent to the remote expert. Both users can see the annotations in the live video.

5. Best practice solutions

While implementing a simple remote support prototype is quite straight-forward, we identified some stumbling blocks for generating a stable and fully operational application. Some technical issues arise from the current state of 3rd-party hardware and software, where we limited ourselves to free software. User-related issues were identified together with company partners by evaluating multiple iterations of the remote support system. In this section, we present our best practice solutions for implementing an open-source remote support system based on off-the-shelf hardware and software components.

Configurability — The optimal configuration of an AR remote support tool depends on environmental restrictions and a wide range of different use cases. Environmental restrictions are the hardware limitations of the portable device, the communication bandwidth, and the 3D reconstruction for the AR view. A good configuration can be automatically determined for some of these restrictions, but sometimes manual adaptions are useful as well (see issue Performance, Stable anchoring and Bandwidth).

Depending on the use case, different communication forms are advantageous. Assembling or dismantling of parts requires different supporting options from documentation of tasks that have to be repeated several times. Hand movements and processes must also be visualized differently than a simple selection of parts. Therefore different tools and communication settings are required. We offer a visualisation of the mouse cursor versus a drawing overlay (see Figure 3) or a permanent anchoring versus a temporary hint (see Figure 5). The tool selection cannot be made automatically based on the use case. Also, the optimal support tool is not the same throughout the session. There are often a number of steps within a maintenance call that require different support. To keep the client of the on-site worker simple, the remote expert manages the optimal tool selection manually.

Bandwidth — A good internet bandwidth cannot be guaranteed in all maintenance scenarios, such as in tunnel construction or industrial plants. Since it is not possible to reduce the video resolution at runtime with our chosen WebRTC Unity plugin, only the frame rate can be adjusted automatically. In case that this is not enough, we offer a special low bandwidth mode [14], which can be activated by the remote expert. In this mode, the transmission of the live video is stopped. The on-site worker determines which images are sent to the remote expert as still frames. The remote expert can then annotate these images as described earlier.

Transmission delay — Our chosen WebRTC Unity plugin offers a dedicated data channel, but the packet size for a message is limited. It is therefore necessary to transmit extensive annotations or high-resolution still frames in blocks. This decomposition can lead to a delay, which is particularly noticeable in the gallery when annotations and still frames are requested from the on-site worker's device. Therefore, although the data is organized by the on-site worker application, it is additionally buffered and synchronized by the remote expert device, so that this delay only occurs when the annotations created by the on-site worker are retrieved for the first time[5]. Because of this delay, the data channel must be relieved as much as possible. Shifting performance-intensive tasks such as the calculation of 3D space to the powerful remote expert PC [19] are therefore not possible.

Security — Security concerns must be taken seriously, because internal company data may be transmitted during a remote support call. By using WebRTC, the devices of the remote expert and the on-site worker communicate directly with each other. This peer-to-peer connection for data transmission and the hosting of the server for the connection establishment (signaling, STUN and TURN server) in the company's own network counteract security concerns. All relevant services are stored within the company and can be managed by the IT security department [4]. Additional security-related aspects, such as the preferred encryption standards or the handling of data and rights management, can also be determined by the company's own IT security. Strong encryption technologies can protect the transferred data. Company or customer data may not be stored on the portable devices. Security warnings and non-disclosure agreements might be necessary, when remote support is executed at customer facilities as video capturing tools may be prohibited.

2D Space — AR offers the possibility of anchoring an annotation in the tracked 3D space and displaying it in the live video. This enables a better assignment of the annotation to the real object by viewing the annotation from different perspectives. However, the mere display of the annotations in the AR space sometimes leads to problems (see issue Performance, Hands-free solutions and Stable anchoring). Therefore, in addition to the 3D space, we have also implemented a 2D space. All annotations are created in 2D space by annotating a 2D screenshot of the live video with a 2D drawing. The transfer into the 3D space takes place by projection. All annotations are displayed in a gallery, which documents the course of the conversation. By selecting a gallery item, the original screenshot is displayed and overlaid with the 2D annotation (see Figure 4). Both users see the annotation from the original perspective, which facilitates the interpretation of unclear projections in 3D AR space (see issue Stable anchoring).

Annotation permanence — In terms of time, the decisive factor is how long annotations are relevant. If real environment changes with every work step, e.g. due to assembling, permanent anchoring of all hints would fill up view and cause confusion (see Figure 5 left). We offer two drawing modes: live drawing and permanent drawing. With the permanent drawing mode, the annotation is permanently stored in the 3D scene and in the gallery. This is useful for labeling parts or repetitive work such as selecting a button when operating a machine. The

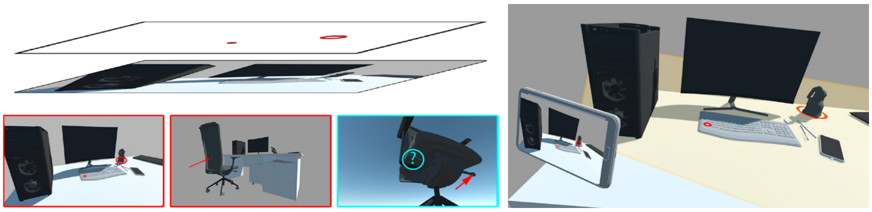

Fig. 4: Comparison of 2D (left) and 3D (right) annotations. 2D annotations are displayed in the gallery as 2D screenshots overlaid with a 2D drawing. 3D annotations are displayed as AR content in the live video.

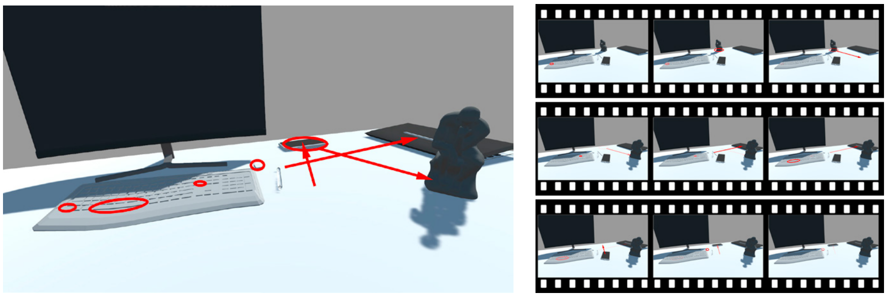

Fig. 5: Permanent anchoring (left) vs. a temporary hint (right). Permanent anchors are displayed for the whole session. If the real content is moved, the annotation remains at the old location, which can fill up the view and cause confusion. Temporary hints disappear automatically after a few seconds.

annotation needs to be meaningful on its own without accompanying voice instructions. Therefore, the annotation is only sent to the on-site worker after the remote expert has finished it. Annotations in the live drawing mode automatically disappear after a short time (see Figure 5 right). The annotations are immediately transferred, i.e. every new brush stroke is visible to the on-site worker. This mode enables quick communication for individual steps. The remote expert can quickly toggle between permanent and temporary drawing depending on the current use case.

Performance — Tracking and reconstructing the 3D space with ARCore or ARKit is performance-intensive, which quickly overheats older smartphones and drains the battery. Remote support consultations are usually longer sessions and do not require permanent AR support. To extend the possible session duration, the remote experts can temporarily switch off the AR calculation. By deactivating the 3D reconstruction, the video signal is tapped directly from the camera without any further calculations. Annotations can be created as described, but they do not appear in the live video in any device. The annotations are only visible in the gallery in the 2D view (see issue 2D Space). In non-AR mode, whenever a new annotation is created, the application automatically switches to the gallery view to make other users aware of the new annotation. For details on the gallery view, see section 3. The remote expert can switch between AR and non-AR mode at any time. If AR is switched back on, previously created annotations are visible again at the correct location. However, annotations created during the non-AR mode are never visible in the live video, since no 3D positions were calculated.

Hands-free solutions — In order to view the anchors in 3D space, the smartphone must be held constantly by the on-site worker. Here, too, the 2D gallery space can be helpful. The snapshot with the annotation is retained in the gallery even without the camera being aligned with the scene. This allows the on-site worker to put down the smartphone and use both hands for the work and at the same time see the remote expert's information (see Figure 6). When the smartphone is put down, the AR camera no longer records relevant video information. It is therefore advisable to temporarily deactivate AR calculation (see issue Performance) and live video transmission (see issue Bandwidth). These measures save performance and prevent a useless video image from being transmitted to the remote experts.

Fig. 6: 2D vs 3D handling of the smartphone.

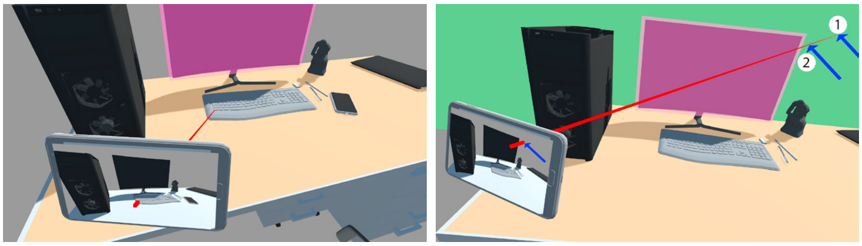

Fig. 7: Left: The destination projection plane is automatically determined by sending a raycast from the click point to the reconstructed 3D scene. Right: The automatic plane selection triggers the wrong plane (green wall plane). With the manual plane correction the user can choose between all found AR planes within the current camera field of view to manually adjust the desired projection plane (pink screen plane).

Smart glasses guarantee hands-free work, which made them highly recommended for our company partner at the beginning of the research project. Therefore we also support Android smart glasses as on-site worker devices. However, their relevance has gradually decreased in the course of development. The devices available on the market are usually not suitable for industrial use in terms of robustness, battery life and ergonomics. Monocular devices require a permanent refocusing of the eye between reality and AR content. In addition to the technical limitations, market penetration is an essential factor in remote support. On-site workers often carry smartphones with them but not smart glasses. In addition, most Android smart glasses do not support ARCore. In this case annotations are only displayed in the 2D space. The operation of smart glasses requires a dedicated GUI concept. Text input for the remote expert's email address or the change between the views and the gallery entries are difficult to achieve. Therefore, we provide a simplified function set for smart glasses support, in which the remote expert also controls the view of the on-site worker. We also tested Microsoft Hololens as an end device for on-site workers. While it offers optimal conditions for remote support such as the 3D reconstruction, the integration with other systems is difficult. Many parts of the system have to be implemented separately, such as the connection via WebRTC, the positioning of AR anchors and the user interface. General support for a wide variety of smart glasses is therefore only possible with great development effort.

Stable anchoring — AR algorithms enable real space to be expanded with additional digital information. The annotation is created in the 2D space of the frozen live image and needs to be transferred to a 3D anchor.

We use two possibilities for transferring the 2D annotation to 3D: dominant planes or planes parallel to the camera [9]. A raycast is sent from the click point to the 3D scene (see Figure 7). If the raycast hits a reconstructed plane, the annotation is projected onto this plane. The annotation is correctly visualized from many viewpoints provided that the object of interest is located on the reconstructed plane. If the raycast hits a feature point, the annotation is anchored to this point and projected onto a plane parallel to the camera. The scene is usually not located at the same depth as the click point all over the image. Therefore, annotations on planes parallel to the camera may be overlaid to very different objects when they are viewed from different angles (see Figure 8).

As shown in Figure7, the automatic plane selection leads to a suboptimal projection, when the annotation is drawn in the free space next to the object of interest. In this case, the remote expert can change the projection plane manually during the drawing activity (see Figure 9).

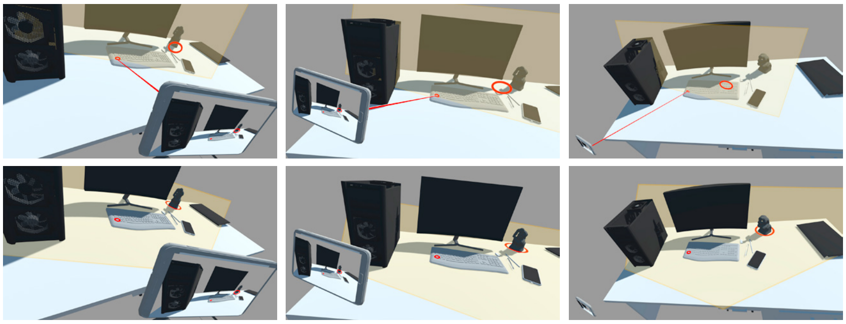

Fig. 8: Anchoring in the click point (top) vs projection on a detected plane (bottom).

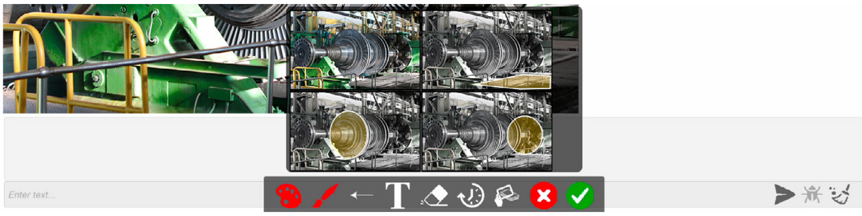

Fig. 9: The remote expert can change the projection plane by selecting the preview image.

Hardware restrictions — As the AR space is calculated on the smartphone, the hardware support of the on-site worker determines the available anchoring options. Older smartphones sometimes do not offer ARCore or ARKit support, which is the basic technology for the AR calculation used. Anchoring in 3D space is not available with these devices and the system switches to non-AR mode (see issue Performance).

Environmental conditions — Ambient noise, lighting conditions and internet bandwidth influence the quality of AR remote support. Industrial environments can have a very high noise level, which makes communication via telephone difficult. We have integrated a chat function to enable communication in these environments. Voice control of smart glasses is therefore not possible in such environments, which also supports our preference for smartphones and tablets (see issue Hands-free solutions). The internet connection is also very poorly developed in some industrial plants or in tunnel construction sites. To counteract this bandwidth problem, we have introduced a low bandwidth mode (see issue Bandwidth). Lighting conditions also have a major impact on communication. Poor lighting influences the video quality for the evaluation by the remote experts and the quality of the AR tracking. Poor AR tracking can cause the annotations to slip, which promotes misunderstandings. AR offers great potential, but is not yet technically 100% reliable in all devices and lighting conditions. The 2D space is an important fallback solution for failure cases.

6. Evaluation

The previous sections described implementation details as well as best practice solutions. Our current prototype was tested and evaluated in a controlled setting at (anonymized). The test setup and the results of this preliminary evaluation are described in detail in [16]. In the following Section we sum up the test setup and the most important findings, since this information forms the basis for our discussion in Section 6.4.

The goal of the evaluation was to get insights into the prototype's potential in direct comparison to paper-based maintenance instructions (which are still commonly used). Additionally, we wanted to learn, if the prototype is free of bugs and technical "fatal-flaws", which could prevent further realistic evaluations in industrial settings.

6.1. Methodology

For the evaluation, 30 test persons with different technical skill levels were invited to conduct a predefined maintenance procedure. The task was to change a specific heat pipe in an industry-PC. Participants in the experimental group were guided by a remote expert using our remote support application on a smartphone. The control group used traditional paper-based instructions, containing 15 variants of the PC, each of them having different wiring and heat pipes to be changed. We took track of the overall time to complete each step as well as the number of tasks completed incorrectly.

6.2. Study content

When defining the evaluation setup, a realistic, industry-related maintenance task was important for our company partners and for us. The job was divided into the following tasks: (A) Type identification, (B) Cover removal, (C) Heat pipe removal, (D) New heat pipe assembly. The following paragraphs describe these tasks and their challenges.

A: Type identification - The large number of machine and equipment types is a challenge for the maintainers. Regarding paper-based instructions, the participants had to compare the cable pattern with 15 pictures in the first chapter of the manual. If they chose the wrong variant, the following steps could not be conducted correctly. With the AR remote support tool, the participants positioned the camera in a way that the remote expert could identify the correct variant.

B: Cover removal - To remove the cover, two screws had to be loosened. Afterwards, the cover had to be moved along a predefined movement. This was the most complex movement in the maintenance process. The participants had to find the spot where to push and in which direction they had to move the cover. This turned out to be difficult, especially on the basis of static pictures. With AR remote support, the remote expert was able to direct the on-site worker into the right position more clearly and correct possible mistakes immediately.

C: Heat pipe removal - For this task, one specific heat pipe had to be changed. Compared to the previous task, the heat pipe removal was less complex. In the paper-based instructions, a picture of the opened PC is shown and the relevant screws are marked with bold red circles. A written task description guided the on-site workers through the process. With AR remote support tool, the remote expert additionally checked the removal steps and corrected occurring mistakes.

D: New heat pipe assembly - In the last step, the new heat pipe had to be placed correctly and mounted with aluminium plates and four screws. Depending on the PC variant, the screws had to be placed as specified. Finally, the cover was mounted.

6.3. Study results

It took both groups nearly the same amount of time to fulfill the maintenance task: 4 minutes and 55 seconds with paper-based instructions, 4 minutes and 57 seconds with AR remote support tool. The situation is different when looking at the overall error rate: 53% of the PC maintenance tasks were carried out erroneously with the use of paper-based instructions. Only 13% of the tasks were performed incorrectly with the use of AR remote support. Since there was no significant difference between the two groups in terms of total duration, we looked at the individual tasks A to D separately. Here we see that the duration for task A with the use of AR remote support tool is significantly shorter compared to paper-based support. Task B also has a shorter duration with AR remote support, but the time advantage in this case is only minimal. The opposite is the case for task C - D, where the paper instruction shows a slight time advantage. For more details on the study and its results, see [16].

6.4. Discussion and further applications

The analysis of the experiment results shows that the main potential of AR remote support is the reduction of errors in maintenance activities. The decisive factor is the minimization of errors in critical tasks. The first step (A) influences all further steps. If the wrong variant is selected, all subsequent tasks are incorrect. Remote expert consolidation with the help of an AR remote support tool therefore has a great advantage for critical steps. The total duration for the maintenance processes was similar with both tested methods. Additionally, time for the connection establishment and the working time of the remote experts have to be included. Nevertheless, we have seen that remote support is significantly faster when machine components, types or variants have to be identified.

Steps A-D have shown that there are differences in the usefulness of AR remote support. Therefore, a broad feature set is beneficial to support different use cases. Augmented reality is not necessary for the first step (A), the video transmission is sufficient for the type selection by the remote experts. In step B, rapid communication and temporary hints are the best option, since the maintenance work changes the real environment. Visualizing the remote expert's mouse cursor in the AR scene is useful for finding the required buttons (see Figure 3). Also the temporary live drawing is a useful option for drawing the movement directions in step B. AR has an advantage for this step compared to 2D space, because the on-site worker can directly identify the right buttons in the live camera image from any position in the room. For steps C and D, permanent annotations within the live video are not useful either, since these are also (dis-)assembly steps that change the real environment. In the case of C and D, 3D anchoring on handheld devices hinders the execution of the main task by tying up the hands of the on-site worker. Using the 2D space is recommended, as this enables the smartphone to be put down. In 2D space the on-site worker sees the screenshot with the annotation and has the hands free for the maintenance task (see issue Hands-free solutions).

Permanent AR annotations within the live video were never an advantage in the maintenance case described. Permanent AR anchoring would be beneficial for repetitive tasks and documentation for future work. However, since our test referred to a short maintenance activity, the advantage of permanent AR anchoring could not be tested. The evaluation has shown that a flexible tool selection is necessary for supporting different use cases.

7. Conclusions and future directions

The cooperation with our company partners revealed great interest in using mobile remote collaboration for maintenance work. For the companies, an in-house solution of remote support has the benefit that it can be well-configured based on the users and the expected use cases. Furthermore it allows the integration into existing processes and enhanced security options. In this paper, we showed how a remote support system can be implemented based on standard software and hardware components. We described the system architecture, underlying software components and user interactions, which we optimized together with industrial partners. Furthermore, we identified general issues when implementing remote support and offer best-practice solutions. These include low internet connection, hands-free solutions, security concerns, performance issues and the stability of annotations. Future remote support applications can be better prepared by knowing these challenges in advance. An interesting finding was, that solutions without augmented reality are an important extension and can solve multiple problems. In general, we think that a broad set of features which can be turned off and on are beneficial. Finally, we compared the AR remote support system with paper-based maintenance instructions. The results were promising and showed a reduced number of errors when AR remote support was used. By taking a closer look, we observed that the individual tasks in the user study required different tools. This confirmed our assumption that a remote support system should contain a configurable set of features. To prove this assumption, another user study is planned to investigate the preference of different configurations for different use cases, e.g. AR vs. non-AR mode or 2D annotation vs. pointer.

Our future work also includes the integration into existing processes and software units. Companies already use ticketing systems and CMMS. The occurrence and results of remote support sessions need to be captured in the CMMS and possibly start other processes. It is also beneficial to include information from other sources into the remote support system. For example, the remote expert can directly link the on-site worker to existing instructional videos or manuals. In some cases, live remote support is not possible due to a lack of internet connection or because remote experts are not available all the time. Asynchronous methods allow sharing of spatial information although on-site workers and remote experts are not connected at the same time [17]. Ideally, it is possible to switch seamlessly between live and asynchronous connections. Furthermore, it would be beneficial to convert remote support sessions into documentation. In case of reoccurring problems, the on-site workers may find a solution in the available documentation in the first place or the remote expert can refer to it. AR remote support is an intensively investigated topic in the scope of digitalization. Even if the technology still has to evolve, the pace of innovation picks up and industry will soon be able to exploit the technology's full potential.

Acknowledgements

This work was funded by the Austrian Research Promotion Agency (FFG) under the project Mixed Reality Based Collaboration 4 Industry (FFG Collective Research).

References

[1] Bordegoni, M., Ferrise, F., Carrabba, E., Di Donato, M., Fiorentino, M., Uva, A.E., 2014. An application based on augmented reality and mobile technology to support remote maintenance, in: Conference and Exhibition of the European Association of Virtual and Augmented Reality, pp. 131–135. doi:10.2312/eurovr.20141351.

[2] Chen, S., Chen, M., Kunz, A., Yantaç, A.E., Bergmark, M., Sundin, A., Fjeld, M., 2013. Semarbeta: mobile sketch-gesture-video remote support for car drivers, in: Proceedings of the 4th Augmented Human International Conference, pp. 69–76. doi:https://doi.org/10.1145/2459236.2459249.

[3] Christiansen, B., 2018. Cmms & maintenance statistics you should know going into 2018. https://limblecmms.com/blog/cmms-maintenance-statistics-for-2018/. Accessed: 2020-04-27.

[4] Communications, N., . A study of webrtc security. http://webrtc-security.github.io/. Accessed: 2020-05-05.

[5] Domova, V., Vartiainen, E., Englund, M., 2014. Designing a remote video collaboration system for industrial settings, in: Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces, pp. 229–238. doi:https://doi.org/10.1145/2669485.2669517.

[6] Fakourfar, O., Ta, K., Tang, R., Bateman, S., Tang, A., 2016. Stabilized annotations for mobile remote assistance, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 1548–1560. doi:https://doi.org/10.1145/2858036.2858171.

[7] Gauglitz, S., Lee, C., Turk, M., Höllerer, T., 2012. Integrating the physical environment into mobile remote collaboration, in: Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services, pp. 241–250. doi:https://doi.org/10.1145/2371574.2371610.

[8] Gauglitz, S., Nuernberger, B., Turk, M., Höllerer, T., 2014a. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation, in: Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology, pp. 197–205. doi:https://doi.org/10.1145/2671015.2671016.

[9] Gauglitz, S., Nuernberger, B., Turk, M., Höllerer, T., 2014b. World-stabilized annotations and virtual scene navigation for remote collaboration, in: Proceedings of the 27th annual ACM symposium on User interface software and technology, pp. 449–459. doi:https://doi.org/10.1145/2642918.2647372.

[10] Gurevich, P., Lanir, J., Cohen, B., Stone, R., 2012. Teleadvisor: a versatile augmented reality tool for remote assistance, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 619–622. doi:https://doi.org/10.1145/2207676.2207763.

[11] Kim, S., Lee, G., Sakata, N., Billinghurst, M., 2014. Improving co-presence with augmented visual communication cues for sharing experience through video conference, in: 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE. pp. 83–92. doi:https://doi.org/10.1109/ISMAR.2014.6948412.

[12] Kim, S., Lee, G.A., Ha, S., Sakata, N., Billinghurst, M., 2015. Automatically freezing live video for annotation during remote collaboration, in: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1669–1674. doi:https://doi.org/10.1145/2702613.2732838.

[13] Loughran, M., 2017. The top maintenance statistics you should know for 2017. https://blog.capterra.com/the-top-maintenance-statistics-you-should-know-for-2017/. Accessed: 2020-04-27.

[14] Masoni, R., Ferrise, F., Bordegoni, M., Gattullo, M., Uva, A.E., Fiorentino, M., Carrabba, E., Di Donato, M., 2017. Supporting remote maintenance in industry 4.0 through augmented reality. Procedia manufacturing 11, 1296–1302. doi:https://doi.org/10.1016/j.promfg.2017.07.257.

[15] Nuernberger, B., Lien, K.C., Höllerer, T., Turk, M., 2016. Interpreting 2d gesture annotations in 3d augmented reality, in: 2016 IEEE Symposium on 3D User Interfaces (3DUI), IEEE. pp. 149–158. doi:https://doi.org/10.1109/3DUI.2016.7460046.

[16] Obermair, F., Althaler, J., Seiler, U., Zeilinger, P., Lechner, A., Pfaffeneder, L., Richter, M., Wolfartsberger, J., 2020. Maintenance with augmented reality remote support in comparison to paper-based instructions: Experiment and analysis, in: 7th International Conference on Industrial Engineering and Applications (ICIEA 2020), IEEE, Bangkok, Thailand (online event). pp. 942–947. doi:https://doi.org/10.1109/ICIEA49774.2020.9102078.

[17] Reisner-Kollmann, I., Aschauer, A., 2020. Design and implementation of asynchronous remote support, in: Proceedings of XChange Reality! 2020, St. Pölten, Austria. pp. 12–15.

[18] Rice, M., Ma, K.T., Tay, H.H., Kaliappan, J., Koh, W.L., Tan, W.P., Ng, J., 2018. Evaluating an augmented remote assistance platform to support industrial applications, in: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), IEEE. pp. 592–597. doi:10.1109/WF-IoT.2018.8355133.

[19] Schneider, M., Rambach, J., Stricker, D., 2017. Augmented reality based on edge computing using the example of remote live support, in: 2017 IEEE International Conference on Industrial Technology (ICIT), IEEE, Toronto, Canada. pp. 1277–1282. doi:https://doi.org/10.1109/ICIT.2017.7915547.