Interacting and Cooperating Beyond Space:
Tele-maintenance within a Virtual Visual Space

Michael Kleiber, Carsten Winkelholz, Thomas Alexander,
Frank O. Flemisch, Christopher M. Schlick
Fraunhofer FKIE
Fraunhofer-Str. 20, 53343 Wachtberg
GERMANY

michael.kleiber@fkie.fraunhofer.de

ABSTRACT

This paper describes an innovative concept and implementation of a maintenance system enabling tele-cooperation of distributed technical personnel. It provides a synchronous shared visual workspace in remote locations with limited bandwidth. A mechanic at the remote location uses an Augmented Reality (AR) system which is connected to the Virtual Reality (VR) system of an expert at the homeland. The expert interactively creates 3D instructions on his VR system that are displayed on a ruggedized hand-held tablet computer of the mechanic. The mechanic considers and follows these instructions during his maintenance work. In addition he can interact in real time with the live AR view creating spatial references for the expert.

The system has been evaluated by 18 experienced automobile mechanics, 6 of them technical soldiers of the German army. The maintenance task consisted of the disassembly of the camshaft housing of an internal combustion engine. The results show clearly that participants completed significantly more tasks and used less verbal instructions when using the VR system compared to a video system. Thus, performance was increased by the system. In the paper we will detail the concept, give an overview of the implemented system and present the results of the practical evaluation.

1 INTRODUCTION

Mechanical engineering and construction is the leading industrial sector in Germany (Wiechers and Schneider, 2012). Maintenance for technical machines and devices is usually handled by 3rd party companies (80% of all foreign subsidiaries are related to service). As maintenance costs have become an essential component of the total cost of ownership the reliability of the machines is a strong purchase criterion. Reliability can be increased with highly qualitative maintenance.

The importance of fast and qualitative maintenance has also grown within the armed forces because of the diversity and variety of platforms and technical equipment. As it has become impossible to employ maintenance technicians specially trained for each piece of equipment, the on-site technician has to be able to effectively and efficiently cooperate with a remote expert using telecommunication equipment. This approach is subsumed by the term “tele-maintenance” (Sanchez et al., 2011). Sharing a visual space in such a tele-maintenance session may improve communication and interactive cooperation (Fussel et al., 2000). However, the results of tele-maintenance using video communications are still inferior to the results achieved when cooperating directly (Alexander, 2012).

1.1 Maintenance

According to DIN EN 13306 (2012) maintenance describes a “combination of all administrative and technical actions during the life cycle of an item intending to retain it in, or restore it to, a state in which it can perform the required functions”. It includes different types of maintenance activities: preventive maintenance, inspection and corrective maintenance, etc. Corrective maintenance is an activity performed as a result of failures or deficiencies, to restore items to a specific condition. This activity may consist of repair, restoration or replacement of components. The cost associated with corrective maintenance is ca. 30% of the total maintenance cost. As corrective maintenance cannot be planned an expert technician is needed on short notice.

Reducing the costs of overall system maintenance or increasing the quality of maintenance carried out can be achieved by employing administrative actions, e.g. regularly scheduled preventive maintenance or using methodologies such as TPM (Total Productive Maintenance) or TQM (Total-Quality-Management). Another option is the employment of advanced or extended diagnostic and support systems, e.g. connecting the OBD (On-Board-Diagnosis) to an interactive maintenance manual.

Interactive electronic technical documentation (IETD) provides animated views of critical components. Videos of lessons learned for special maintenance procedures will also be included for detailed references. This approach may support a more intuitive or faster way to comprehend repair and maintenance instructions. The electronic documents follow the S1000D standard which also incorporates ways to provide 3D CAD data (S1000D, 2008). The process of creating these electronic documents is supported through tools such that the maintenance and logistic information is incorporated in an electronic database. Producing the electronic or paper documentation can therefore be done based on a single source (Wampler et al., 2002).

1.2 Remote Cooperation

Besides integrating an electronic maintenance manual with the diagnostic unit other functions such as central work management or audio/video communications will be integrated as well. Integrating communication equipment offers new functionalities for remote support of technical maintenance personnel. The special case of supporting maintenance technicians of the German army by linking them to a support center in Germany is shown in Figure 1 as an example.

Figure 1: Maintenance technicians in worldwide remote locations are supported by experts at homeland locations

By remotely cooperating both people involved try to solve the problem at hand. The subtasks of the process are: problem description, problem comprehension, task description, task comprehension and finally task execution. The maintenance problem will not be solved in a single iteration of this cooperative cycle but will require a sequence of cycles as depicted in Figure 2.

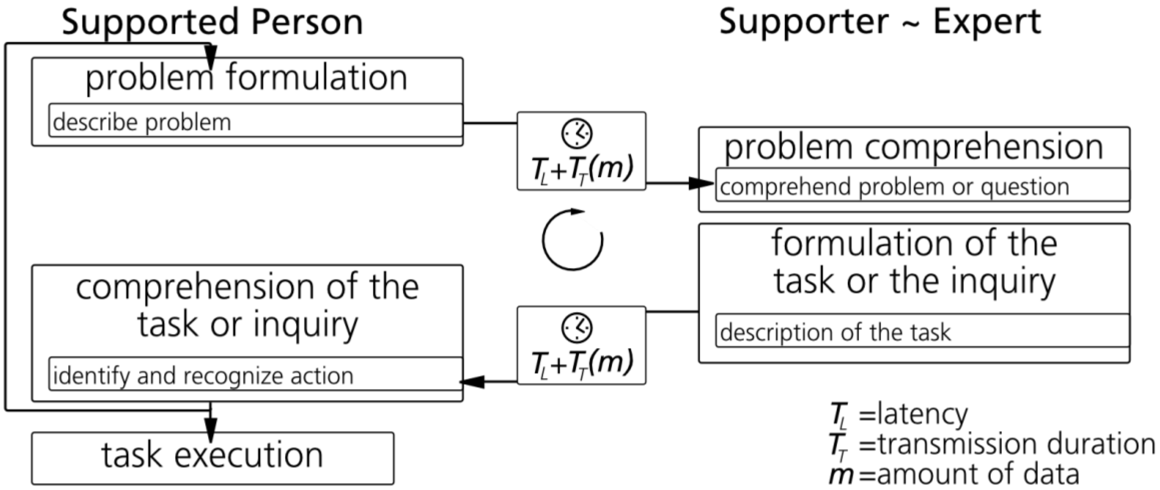

Figure 2: Remote cooperation for support

1.4 Use of Augmented and Virtual Reality

Technologies for audio-visual remote cooperation have been available for decades already. Despite increasing network bandwidth there are still many situations where only limited data exchange is possible. Yet, powerful mobile devices and data storage allow a technician to access a comprehensive database of support information by means of a portable computer system. It can also provide step-by-step visual instructions. By utilizing AR-technologies additional visual information about an assembly or a maintenance procedure can be integrated into the real scene (Barfield et al., 2001). Existing systems were limited to standardized, well-known procedural tasks. By utilizing a remote expert the technician can attain further support.

Applying AR has frequently been documented as support for industrial purposes. There have been large collaborative research consortiums such as ARVIKA (Friedrich, 2002), STAR (Raczynski and Gussmann, 2004) and ARTESAS (Haberland et al., 2007) which have investigated the application of AR for almost all aspects of manufacturing. The according research activities describe rather elaborated system concepts which involve multiple cameras, multiple computers and are in general quite complex.

Utilizing AR for maintenance tasks has slightly different requirements. Henderson and Feiner (2009) categorize maintenance as consisting of activities involving the inspection, testing, servicing, alignment, installation, removal, assembly, repair, overhaul, or rebuilding of human made systems. In these categories, assembly tasks have received the most attention. These tasks can range from assembling aircraft wire bundles (Curtis et al., 1999) to assembling medical equipment for minimal invasive surgery (Nilsson and Johansson, 2007).

However, in most of the related work tele-cooperation is of only minor importance. The AR applications rather resemble an extended electronic handbook which the technician uses without additional help by an expert. The user-friendly creation of AR scenes consisting of a set of maintenance instructions is therefore an important topic in the work of the research consortia cited above as well as in other projects (Knopfle, 2005).

2 SYSTEM CONCEPT AND IMPLEMENTATION

2.1 Integrated Augmented and Virtual Reality Remote Maintenance

Interactive cooperation between an expert at home location and a remote local technician requires sufficient bandwidth for a synchronous transfer of audio-visual information within the network. A lack of visual information requires the technician to describe potential visible malfunctions verbally. It also requires the expert to guide and direct the technician by means of verbal descriptions. However, sharing a visual space through the use of video improves communication and interactive cooperation (Fussel, 2000). Our approach to bridge the gap between the need for visual information exchange and a narrowband network is based on a virtual reconstruction of the maintenance object through the use of AR techniques.

The expert uses an egocentric virtual view derived from the mechanic's view. This allows the expert to formulate instructions within a spatial context and it also gives the questions of the mechanic a spatial frame of reference. Thereby, expert and mechanic share a visual space which is beneficial in collaborative physical tasks (Gergle, 2005). Instead of transmitting a video our system first identifies the machine parts in the view and then transmits the IDs of these parts as well as their location and orientation in a specific coordinate system. The expert uses a VR system which reconstructs a 3D view of the mechanic's point-of-view (Figure 3). A precondition for this concept is that 3D models of the maintenance objects and their subparts are available and the real objects can be identified.

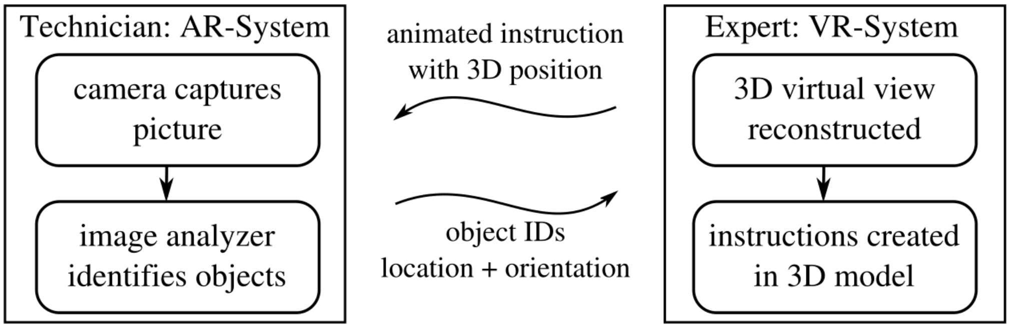

Figure 3:The concept underlying the integrated AR-VR tele-maintenance system.

In our case, the expert uses a desktop VR system to view and interact with the virtual 3D view. Besides the virtual view from the mechanic the expert can examine an interactive model of the maintenance object from an arbitrary viewpoint. Thereby he is able to explore the object interactively, e.g. to plan the repair of a machine. Other functions of the interactive VR system include adding text annotations at 3D positions, creating animated 3D instructions or placing visual hints at 3D locations.

For the AR system of the mechanic we first developed a concept with a head-worn display. However, the weak acceptance of these displays by maintenance personnel because of their limited field of view and considerable weight led to a concept revision. Instead, a tablet computer with a backside camera is used. This paper is focused on the VR application of the expert. A detailed description of the AR system and the results of an evaluation were published earlier (Kleiber, 2011).

2.3 Implementation

The hardware of the VR system consists of two stereoscopic screens, one for the interactive and one for the passive 3D view (Figure 4). We have decided to use stereoscopic displays for the work place of the expert because we believe that the spatial presence of the expert benefits from the additional depth cue of disparity (Schlick, 2011). This is especially important for the reconstructed passive viewpoint of the mechanic. Whenever the mechanic moves the camera the expert has to reorient and identify which parts are currently in the view. A stereoscopic 3D view allows a quicker and more reliable orientation and localization after a change in position or orientation of the camera (Kleiber, 2012).

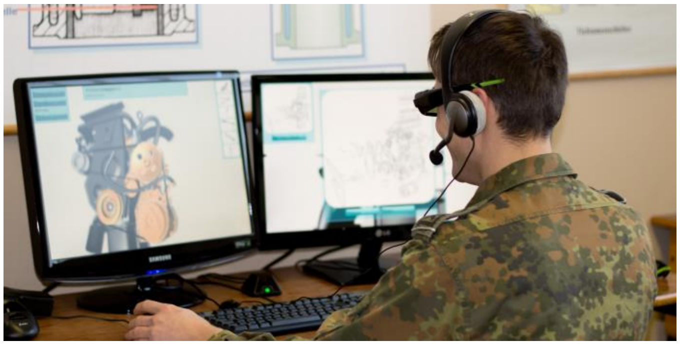

Figure 4: A participant using the two screen system wearing shutter glasses.

As a matter of fact we have also decided to use 120 Hz LCDs with shutter glasses instead of line polarized LCDs because of the better image quality. The computer driving these displays is a standard PC with a quad buffered graphics card and achieves 30 frames per second and per eye. As input devices we use a standard 2D mouse for selection and menu interaction, a keyboard for text entry and a Logitech 3D mouse for navigation.

To reduce eyestrain and adaption times we have paid special attention to the generation of comfortable stereoscopic images which can be perceived immediately when looking at the stereoscopic displays. This is important since visual discomfort can strongly impact the usability of the system (Lambooij, 2007). To achieve this stereoscopic projection parameters like location of image plane and stereo base were adapted depending on the camera position within the virtual environment.

We have adapted a technique developed for 3D object inspection (Kleiber, 2009). The user selects a point of interest by left-clicking with the 2D mouse. The point of selection will become the new point of zero disparity. In a smooth transition the virtual camera is reoriented so that the point of zero disparity is at the center of the screen. The transition is based on a finite impulse response filter with $0.5\,(1 - \cos \pi x)$ as the core function. The amount of stereopsis is based on the distance of the camera position to the point of selection. The field of view could not be changed to increase the comfort of the stereoscopic images, because it was determined by the real camera employed by the mechanic. Using the field of view of the real camera allows the replication of the view of the mechanic. Furthermore, it allows the expert to view a photo taken by the mechanic overlaid over the 3D object. Therefore, the distance of the projection plane is adjusted in order to control the amount of disparity.

Besides adapting the amount of disparity we also use a shader-based real-time depth of field effect to further improve the quality of the stereoscopic visualization. We also place the plane of focus at the point of zero disparity. The amount of defocus increases linearly. The implementation is based on the one by Riguer et al. (2004). Although the effect adds complexity to the rendering pipeline and therefore lowers the refresh rate of the visualization the effect achieved is worth the performance decrease. Especially window violations are not as distracting as when they are in full focus.

On the active 3D view the user can select parts of the maintenance object using the 2D mouse. The selected part can be combined with a tool and an action selected from a menu to create an animated work instruction. It can be put into an instruction package by using a button on the 3D mouse. By combining multiple instructions into a package complex work sequences can be created.

Furthermore, the expert can select any location on the 3D object to position a pointer which is duplicated in the mechanic's view. Precise spatial references are therefore easy to create. Likewise, the mechanic can point at the video view to position his shared pointer. Additionally, text annotations can be created at the positions of the shared pointers.

When there is sufficient bandwidth a live video stream can be send from the mechanic to the expert. The stream can be shown integrated into the 3D scene so that it overlays the 3D visualization. The mechanic can also transmit photos of the maintenance object which can be shown integrated into the 3D view or placed in a photo queue. These photos, as well as existing construction drawings, can be visually annotated and sent back.

4 SYSTEM EVALUATION

4.1 Hypotheses and independent variables

Professional repair and maintenance manuals are usually formulated by technical authors. They are often created without severe time and cost constraints (Wampler, 2002). Current systems for the creation of manuals provide the user access to 3D models in a part database (Cortona3D, 2012). However, the interactive real-time creation of work instructions for an interactive 3D AR application by an expert is a novelty. We were therefore interested in whether an experienced automobile mechanic, i.e. an expert mechanic, would be able to intuitively work with our stereoscopic VR system, e.g. to guide a novice in executing an engine repair task.

The hypotheses are therefore formulated in regard to the efficiency, effectiveness and usability of a 3D system as a tool to interactively give guidance and support in a telecooperation task. The alternative system was considered to be an off the shelf video conferencing tool which allows, beside exchanging audio and video, the graphical annotation of still images. We formulated the following hypotheses comparing our AR-VR system ($S_{\text{AR-VR}}$) with a video based system (S_{Video}):

H_1 The overall time needed for formulating task descriptions or questions is lower when using $S_{\text{AR-VR}}$ compared to S_{Video}.

H_2 The length of textual or phonetic descriptions is shorter when using $S_{\text{AR-VR}}$.

H_3 Visual fatigue will not be significantly higher after using $S_{\text{AR-VR}}$.

H_4 The subjective workload rating will not be higher after using $S_{\text{AR-VR}}$ compared to S_{Video}.

Besides the system used, the introduction of an additional independent variable (system first used) was mandated by our experimental design (see section 4.6).

4.2 Scenario

Tele-maintenance is required when an undocumented maintenance problem arises and bringing in an expert is too costly or time consuming. Communication between the local mechanic and the remote engineer is then usually done using a satellite link. We additionally defined the following boundary conditions in the experimental scenario which are based on the practical requirements of our target user group:

- there is an undocumented defect for which no sequential maintenance procedure exists
- digital technical drawings of the machine are available
- an audio link with an average latency of 2 s is available
- an additional data link with a GSM comparable data rate (14.4 kBit/s) is available with the same latency

Since the technical drawings do not include sequential instructions the experts are required to know the maintenance procedure very well.

4.3 Apparatus

For the expert's work place we used the hardware setup described in section 2.3.

Since the expert's work place was evaluated the mechanic's actions were only simulated by an experiment aide. All sensor data, e.g. photos or camera position, were therefore created beforehand or supplied by the experiment aide. The simulated AR work place consisted of only one monoscopic LCD. A 3D mouse was employed as well.

Both systems were connected with low latency and high bandwidth. The 2 s latency and small bandwidth of the scenario requirements were therefore simulated by restricting the allowed amount of network traffic. Audio communication was implemented using head phones with attached microphones. Audio information and data was transmitted using the same network connection and under the same latency and bandwidth limitations. During the experiment the expert and the simulated mechanic were located in separate rooms. They could therefore only listen to each other by means of an audio connection.

The experiment aide had control over certain functions of the participant’s work place. He triggered the loading of the appropriate 3D model and he activated or deactivated the extended 3D functions.

4.4 Procedure

The goal was to evaluate the implemented system in a real maintenance task of an automobile engine. The exchange of the camshaft housing was taken as it is a complex procedure which consists of 14 work packages with differing number of work steps. Some of the steps may lead to the destruction of the engine. This means that precise instructions are required. The participants of the evaluation had to guide a novice mechanic in the task using the tele-cooperation system under the restrictions outlined in the scenario description. In the experimental evaluation the experiment was finished when the camshaft housing was successfully removed.

We expected large inter-individual differences and therefore chose a within-participant design. We initially planned on interchanging the systems after half of the 14 work packages. Yet, since we also expected that some of the participants would miss some work packages or work steps we reverted to interchanging the system used after 10 minutes. This meant that the participants had 20 minutes to instruct the mechanic. The actual disassembly of the engine to remove the camshaft takes about 1 hour when carried out by an experienced mechanic. Because of this large time difference the participants had to be instructed that the disassembly was simulated. The participants carried out a training using a different maintenance object. They were given small tasks to make sure that they were able to use all of the functions of both systems.

4.5 Participants

Eighteen experienced automobile mechanics took part in the evaluation. They were on average 24 years (SD=4.3 years) old and had on average 5.4 years (SD=3.4 years) experience. All participants had a binocular vision at reading distance of at least 0.7 dpt. The minimum stereoscopic vision acuity was 100". The participants were compensated monetarily.

4.6 Control of Extraneous Variables

The nature of a remote maintenance operation requires that at least two persons are involved. Both persons influence the results. To reduce the extraneous influence the mechanic was played by a single experiment aide. Since the experiment aide has considerable influence on performance and workload we needed to formulate criteria for his behaviour. To document the progress of the maintenance procedure we used the following criteria to decide whether an instruction was considered to be ambiguous. The experiment aide had to inquire further when:

- a part was not or not clearly indicated,
- a task involving a part happened in a new work area, but the expert did not indicate this or did not give a spatial indication about the new location,
- the verbally mentioned part was hidden behind another part,
- the expert used an uncommon term for a part but did not give a description,
- a referred to part existed multiple times but the expert did not indicate this so that a mix-up could occur or
- a wrong part was indicated.

Whereas the second and third of the above events cannot occur when using S_{AR-VR}, producing an unclear or even wrong indication of a part can happen when the AR tracking is inaccurate. To judge whether an indication is exact the experiment aide did a visual validation just like in the real AR system by using pre-recorded photos.

Whenever one of the above events occurred the mechanic asked for clarification by asking: “Where is the specific part located?”", “Can you describe the part in more detail?”, “Can you indicate the part more precisely?” or “Did you mean this part?”.

Besides standardizing the behaviour of the experiment aide we also needed to make sure that all participants had similar knowledge regarding the maintenance procedure. Interviews with experienced automobile mechanics indicated that the chosen procedure is a very common and well known task. Nonetheless, to ensure that all participants knew the engine used and the parts involved, we conducted a preparatory experiment. In this experiment the participants had to locate parts by using the AR system. The parts to be found in the preparatory experiment were the same ones involved in the actual maintenance procedure.

Furthermore, the within-participant experimental design alleviated the problem of individual differences. However, the within-participant design required a change of the systems in the middle of the experiment. As we also expected an influence of the system first used the participants were divided into two additional groups. One group started the maintenance procedure using $S_{AR\text{-}VR}$, the other group first used S_{Video}. Switching the systems used simply meant deactivating the 3D visualization. However, when S_{Video} was used first, the state of the 3D visualization after switching the systems would not reflect the state of the maintenance procedure. In these cases the experiment head adjusted the 3D visualization accordingly. The overall duration of the evaluation was ca. 75 minutes per participant.

4.7 Dependent Variables

Removing the camshaft housing requires a total of 40 work steps when following the recommended procedure. A work step was counted as completed when the participants provided an unambiguous instruction. The number of work steps completed were recorded by the experiment aide. Furthermore, we recorded the number of work instructions, photos and annotated pictures transmitted during the experiment.

The speech of the expert and the mechanic were recorded for later analysis. To evaluate hypothesis H_2 the length of textual and phonetic instructions had to be determined. The analysis was done automatically by using the sound finder tool in the audio editor “Audacity”. The minimum silence duration was set to 1.5 s and the minimum sound duration was set to 0.15~s. These parameters were determined by analysing some of the recordings manually. A quick “ja” (yes) or “OK” took between 0.14 s and 0.20 s.

During the experiment all actions carried out by mechanic and expert were logged with timestamps. This allowed the later explorative analysis of the data, e.g. to calculate individual task durations.

To assess visual fatigue we used a questionnaire based on the one by Bangor (2000). The workload was assessed using the NASA task load index (Hart and Staveland, 1988). We also compared the two systems on a subjective basis using the following three questions which were rated on a scale from 0 (very complicated/bad) to 10 (very simple/good) for both systems used:

Q_1 How simple/hard was instructing the mechanic for you?

Q_2 How good was your spatial conception of the view of the mechanic?

Q_3 How good was your conception of the state of the maintenance object.

4.8 Results

The overall maintenance procedure could have been completed following the accumulated instructions of all mechanics, although no-one provided instructions for the total 40 work steps.

The participants were very motivated and tried their best to give detailed instructions so that even a novice mechanic was able to follow and complete the instructions given. Some of the mechanics reported after the experiment that they were heavily stressed because they wanted to perform well. Some also reported that they had completely forgotten that the maintenance was only simulated.

All but one participant created animated 3D instructions using the 3D functions of the application. However, the one that did not create animated instructions extensively made use of highlighting 3D objects using the shared

3D pointer. This means the training before the actual evaluation was sufficient. Of course, most of the time verbal instructions were given as well. The high latency seemed less problematic than we originally expected.

To evaluate the first hypothesis the number of completed work steps were compared as it indicates how quick generating instructions and answering questions was (see Figure 5). The number of steps were normalized because some participants finished the maintenance procedure before the time was up.

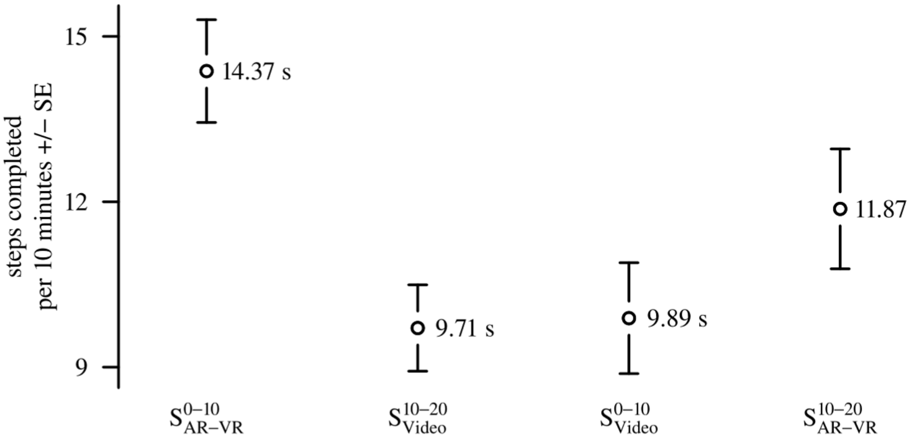

Figure 5: The number of work steps completed. S^{0-10}_{AR-VR} **are the results of the first 10 minutes of using** S_{AR-VR}

A Shapiro-Wilk normality test for the number of steps completed shows no significant deviation from normality for all factor groups. A repeated measures analysis of variance assuming sphericity using the system started with, as a between-participant factor, shows strong significance for the system used ($F_{1,16}$=8.6; p=0.010), but no significance for the starting condition ($F_{1,16}$=2.38; p=0.142). A pairwise comparison of the subgroups using t-tests with p-adjustment according to Holm (1979) shows strong significance (p<0.01) for the one-tailed, paired comparison of S^{0-10}_{AR-VR} with S^{10-20}_{Video}, strong significance (p<0.01) for the unpaired comparison of S^{0-10}_{AR-VR} with S^{10-20}_{AR-VR} and strong significance (p<0.01) for the unpaired comparison of S^{0-10}_{AR-VR} with S^{0-10}_{Video}. The one-tailed, paired comparison of S^{0-10}_{Video} with S^{10-20}_{AR-VR} shows no significance (p=0.3).

None of the participants used textual instructions during the experiment so for H_2 only the verbal exchange has to be considered. The analysis of the sound data showed that the participants talked 4.7% (SD=1.4%) of the time. There is only a negligible difference between the systems. However, the average duration of a talk act of the experts is 2.4 s using S_{AR-VR} and 2.9 s using S_{Video}. Again, a Shapiro-Wilk normality test for the average talk durations of the participants shows no significant deviation from normality. A repeated measures ANOVA using the same factors as above for the duration of talk acts shows strong significance ($F_{1,16}$=21,3; $p < 0.001$) for the system used, but no significance for the starting condition ($F_{1,16}$=1.36; $p = 0.26$). A one-tailed paired t-test shows strong significance (p<0.01) for the comparison of the average talk duration of S^{0-10}_{AR-VR} with S^{10-20}_{Video} and weak significance (p=0.04) for the comparison of S^{0-10}_{Video} with S^{10-20}_{AR-VR}. This means the phrases formulated when using S_{AR-VR} were shorter compared to S_{Video}.

A possible explanation for shorter phrases when using $S_{AR\text{-}VR}$ compared to S_{Video} might be the expressiveness of the visual instructions. When S_{Video} is used visual instructions can only be produced by annotating construction plans or received photos of the maintenance object whereas $S_{AR\text{-}VR}$ allows the creation of animated 3D instructions.

When the participants used $S_{AR\text{-}VR}$ only two participants sent a picture and only one participant requested a photo. However, on average 10.9 instruction packages were transmitted (SD=5.9). When participants used S_{Video} they requested 2.0 (SD=1.1) photos of the maintenance object and sent 5.8 (SD=1.9) annotated photos or construction drawings. This is a clear indication that the participants generated visual instructions more easily using $S_{AR\text{-}VR}$. An unpaired t-test shows that the group starting with $S_{AR\text{-}VR}$ sent significantly ($p<0.01$) more instruction packages than the one starting with $S_{Video.}$

However there is no difference between the groups for the other two measures as can be seen in Figure 6.

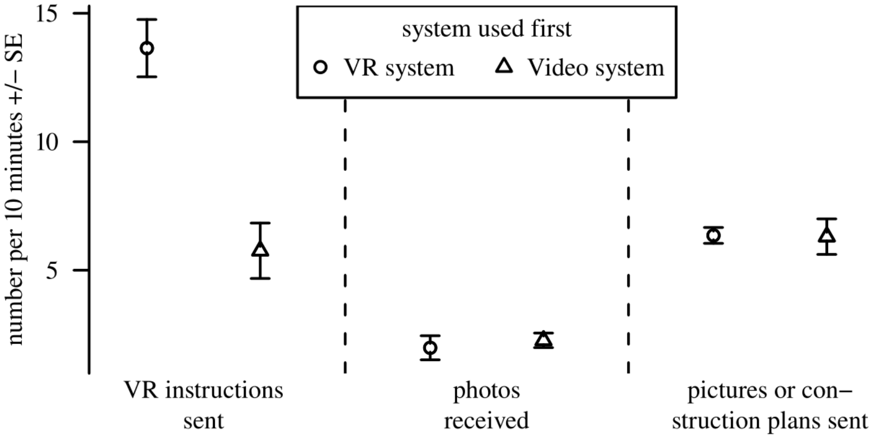

Figure 6: The number of instructions, photos and pictures sent or received.

Hypotheses H_3 and H_4 can only be evaluated using the results of the subjective questionnaires. Since the switch of the system used occurred while the maintenance procedure was on-going we did not assess work load and visual fatigue during the switch but only after the procedure. The participants were therefore instructed to perform an assessment regarding the last system used.

The results of the NASA task load index did not show significant differences between both groups. Yet, it is unclear how well participants were able to exclude their experience with the first system used from their assessment.

The visual fatigue questionnaire did not show significant differences when comparing the data gathered before and after the experiment. A comparison of the two groups also did not show significant differences. This is in accordance with the informal feedback gathered regarding the comfort of the stereoscopic 3D visualization.

4.9 Discussion

Although the participants completed significantly more steps using $S_{AR\text{-}VR}$ the impact of the starting system ($S_{AR\text{-}VR}$ vs. S_{Video}) was considerably large. A similar effect was found for the number of instruction packages sent.

Anecdotal evidence from observations during the evaluation suggests that the difference is likely caused by the sequence of the systems used and not by a difference in the sample groups. After switching from S_{Video} to S_{AR-VR} some participants did have difficulty remembering the exact usage of the 3D part of the application. However, they also did not resort to using the standard teleconference features but rather used a trial and error approach to rediscover the use of the 3D application.

Therefore, a probable explanation for the difference of the groups is that the participants had forgotten the usage of the functions of the 3D application. This was not expected because all participants carried out a training session before the evaluation. The participants also finished training tasks under supervision to review their understanding of the application's functions.

The overall talk duration of the participants did not differ between the systems used. However, since the participants completed more work steps with the same percentage of verbal exchange we must conclude that there is strong significance in favour of hypothesis H_2. Furthermore, this is also supported by the analysis of the length of verbal exchanges. Shorter verbal instructions are also an indication that less descriptions were required. The larger number of instruction packages sent compared to the number of annotated pictures sent and the results of the subjective questionnaire indicate this as well.

Some participants reported that they experienced high cognitive workload. This can be explained by the time constraints of the experiment. The participants were informed about the amount of time available before the experiment. They might therefore have felt pressured since the maintenance procedure in reality takes about 1 hour and only 20 minutes were allotted in the experiment. The time pressure might have also impacted the participants in regard to remembering the use of the 3D application.

Visual fatigue was not observed after the experiment. However, the evaluation during which the participants were required to wear the shutter glasses only took about 40 minutes. Furthermore, any distracting lights which might have caused noticeable flicker were turned off. It is unclear how a prolonged use of a stereoscopic desktop VR system might influence visual fatigue.

5 CONCLUSION AND OUTLOOK

The results show that participants completed more tasks and used less verbal instructions when using the VR system. The subjective evaluation showed a higher rating for the AR-VR system regarding the ease of creating 3D instructions and the mental representation of the state of the maintenance object. Neither an impact on subjective work load nor on visual fatigue were measured. The results therefore support our concept of a stereoscopic 3D system as a beneficial tool for creating instructions in a tele-maintenance task.

Up to now the system was only evaluated for machine maintenance tasks. Yet, the concept is also applicable in the diagnostic phase. Selecting, placing and using the right diagnostic procedure and tools can be supported by a remote expert. Another area well suited for the use of an integrated AR-VR system is training or advanced distant learning. The instructor may use the VR system to instruct multiple students but can also give individual instructions.

REFERENCES

[1] Wiechers, R. and G. Schneider (2012), „Maschinenbau in Zahl und Bild“ VDMA, 2012

[2] Sanchez, C., F. Fernandez, L. S. Vena, J. Carpio and M. Castro (2011), "Industrial telemaintenance: Remote management experience from subway to industrial electronics," Industrial Electronics, IEEE Transactions on, vol. 58, no. 3, pp. 1044–1051.

[3] Fussell, S. R., R. E. Kraut and J. Siegel (2000), "Coordination of communication: Effects of shared visual context on collaborative work," in Proceedings of the ACM conference on Computer supported cooperative work, pp. 21–30.

[4] Alexander, T., C. Pfendler, J. Thun and M. Kleiber (2012), "The influence of the modality of telecooperation on performance and workload," Work: A Journal of Prevention, Assessment and Rehabilitation, vol. 41, no. 0, pp. 3476–3483.

[5] DIN EN 13306:2010-12 (2010), "Instandhaltung - Begriffe der Instandhaltung"

[6] S1000D 4.0:2008-08-01 (2008), "International specification for technical publications utilizing a common source database", S1000D steering committee

[7] Wampler, J., R. Blue, C. R. R. Volpe and P. Rondot (2002), "Service Manual Generation: An automated approach to maintenance manual development", In: Proceedings of the 16th Symposium on Human Factors in Aviation Maintenance, FAA Human Factors, p. 1

[8] Barfield, W., K. Baird, J. Shewchuck, G. Ioannou (2001), "Applications of wearable computers and augmented reality to manufacturing", In: Fundamentals of wearable computers and Augmented Reality, Routledge, pp. 695–713

[9] Fussell, S.R., R.E. Kraut, J. Siegel (2000), "Coordination of communication: Effects of shared visual context on collaborative work", In: Proceedings of the ACM conference on Computer Supported Cooperative Work, ACM, pp. 21–30

[10] Gergle, D. (2005), "The value of shared visual space for collaborative physical tasks", In: CHI extended abstracts on Human factors in computing systems, Portland, OR, USA, p. 1116

[11] Kleiber, M. and T. Alexander (2011), "Evaluation of a Mobile AR Tele-Maintenance System", In: Proceedings of the 6th international conference on Universal access in human-computer interaction: applications and services, Springer Lecture Notes in Computer Science, Volume 6768/2011, p. 253-262

[12] Friedrich, W. (2002), "ARVIKA - Augmented Reality for development, production and service", In: International Symposium on Mixed and Augmented Reality, ISMAR, pp. 3–4

[13] Raczynski, A., P. Gussmann (2004), "Services and training through augmented reality", In: Proceedings of the 1st European conference on Visual Media Production, pp. 263 – 271

[14] Haberland, U., C. Brecher, F. Possel-Dölken (2007), "Advanced Augmented Reality-based service technologies for production systems", In: Proceedings of the International Conference on Smart Machining Systems.

[15] Henderson, S., S. Feiner (2009), "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret", In: 8th International Symposium on Mixed and Augmented Reality, IEEE, pp. 135–144

[16] Curtis, D., D. Mizell, P. Gruenbaum and A. Janin (1999) “Several devils in the details: making an AR application work in the airplane factory”, In: Proceedings of the international workshop on Augmented reality: placing artificial objects in real scenes, A. K. Peters Ltd., Natick, MA, USA, pp. 47–60

[17] Nilsson, S. and B. Johansson (2007), “Fun and usable: Augmented Reality instructions in a hospital setting”, In: Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces, ACM, New York, NY, USA. pp. 123–130

[18] Knopfle, C., J. Weidenhausen, L. Chauvigne and I. Stock (2005), “Template based authoring for AR based service scenarios”, In: Proceedings of the conference Virtual Reality, IEEE, pp. 237–240

[19] Schlick, C. M., M. Ziefle, C. Winkelholz and A. Mertens (2011), “Visual displays”, In: The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications, 3rd ed., ser. Human Factors and Ergonomics, J. A. Jacko and A. Sears, Eds. Lawrence Erlbaum Assoc Inc

[20] Kleiber, M., B. Weltjen and J. Förster (2012), “Stereoscopic desktop VR system for telemaintenance”, In: Proceedings of the conference Stereoscopic Displays and Applications XXIII, vol. 8288, Burlingame, California, USA, SPIE, pp. 828803 – 828803–8

[21] Lambooij; M. T. M., W. A. IJsselsteijn and I. Heynderickx (2007), “Visual discomfort in stereoscopic displays: a review”, In: Proceedings of the conference Stereoscopic Displays and Virtual Reality Systems XIV, vol. 6490, SPIE

[22] Kleiber, M. and C. Winkelholz (2009), “Case study: using a stereoscopic display for mission planning”, In: Proceedings of the conference Stereoscopic Displays and Applications XX, vol. 7237, San Jose, CA, USA, SPIE, pp. 723704 – 723704–9

[23] Riguer, G., N. Tatarchuk and J. Isidoro (2004), “Real-Time depth of field simulation”, In: ShaderX2: Shader Programming Tips and Tricks with Directx 9.0, Texas, USA: Wordware Publishing Inc., 2004, vol. 2

[24] Cortona3D (2012), Cortona3D RapidManual, http://www.cortona3d.com/interactive_electronic_technical_manuals.aspx

[25] Bangor, A. W. (2000), “Display Technology and Ambient Illumination Influences on Visual Fatigue at VDT Workstations”, Dissertation at Virginia Polytechnic Institute and State University

[26] Hart, S. G. and L. E. Staveland (1988), “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research”, Human mental workload Vol. 1, pp. 139 - 183

[27] Holm, S. (1979), “A simple sequentially rejective multiple test procedure”, Scandinavian journal of statistics, JSTOR, pp. 65 - 70