H. Söbke, J. Baalsrud Hauge, M. Wolf, & F. Wehking (eds.):
Proceedings of DELbA 2020
Workshop on Designing and Facilitating Educational Location-based Applications
co-located with the
Fifteenth European Conference on Technology Enhanced Learning (EC-TEL 2020)
Heidelberg, Germany, Online, September 15, 2020

User-Centered Evaluation of the Learning Effects in the Use of a 3D Gesture Control for a Mobile Location-Based Augmented Reality Solution for Maintenance

Moritz Quandt[1], David Hippert[1], Thies Beinke[1],
and Michael Freitag[1,2]

[1] BIBA - Bremer Institut für Produktion und Logistik
at the University of Bremen, Bremen, Germany
[2] Faculty of Production Engineering, University of Bremen, Bremen, Germany
`qua@biba.uni-bremen.de`

Abstract. Mobile Augmented Reality (AR) solutions are ascribed to a high potential for location-based support in the work context. The technology enables the insertion of virtual content directly into the working environment. The successful introduction in practice of the developed solutions is highly dependent on the acceptance of the end-users. Since there are no general design principles for integrating novel forms of interaction and user interfaces into a three-dimensional application environment, we apply user-centered evaluation methods. In this paper, we investigate the learning effects of the users in handling a hand-based gesture control using the example of an AR application to support the maintenance processes of heating, air conditioning, and cooling systems. The users perform five tasks in two successive test runs. Based on the processing times and the required interactions for each task, we can evaluate the applicability of the selected interaction patterns for the respective task.The user study results show that users learn to use hand-based gesture control in a short time. Especially when directly manipulating virtual objects, the users quickly showed improvements regarding processing time and number of interactions needed. In contrast, learning effects in the use of the hand-gesture control do not become evident when performing multi-step gestures without reference to the real environment. Since existing interaction patterns do not necessarily achieve high user acceptance in this context, user studies can provide valuable insights for the design of mobile location-based AR solutions.

Keywords: Augmented reality, location-based information provision, 3D hand gesture control.

---

1 Introduction

The reality-virtuality-continuum of Milgram and Kishino classifies Augmented Reality (AR) as a technology that extends the real world with virtual content. While virtual reality focuses on the complete immersion of the user in a virtual world, AR focuses on the coexistence of real and virtual objects [1]. Azuma defines AR as the combination of virtual reality and the real environment with partial superimposition, interaction in real-time, and a three-dimensional (3D) relationship between virtual and real objects [2]. For work process-integrated support and context-related information provision in the work environment, AR technology is particularly suitable [3], [4]. In the field of industrial applications, product design, manufacturing, assembly, maintenance, and training are seen as the main application areas [5]. In the area of maintenance, several studies have developed promising approaches to improve employees' performance in the execution of technical maintenance tasks, for the training of employees to perform maintenance tasks or to support the documentation of maintenance activities [6]. Especially in maintenance, the documentation and transfer of knowledge of experienced service technicians play an important role. This know-how for the maintenance of machines and components is essential for the efficient processing of maintenance orders. However, only a limited number of AR solutions to support service technicians with location-based information in the work environment have already been used in practice [6].

In order to increase the acceptance of AR-based solutions by users in practice, the optimal interaction between humans and technology is the decisive criterion for the development of AR-based assistance systems [7]. The high number of possible forms of interaction, hardware configurations, and the possibility of addressing different senses (visual, auditory, tactile) obstructs the development of generally applicable, comprehensive design guidelines for AR applications [8]. As a basis for the development of interactive AR systems, general requirements for the design of industrial AR solutions [9] or dialogue principles, e.g. according to DIN 9241-110, are available. The decision on the specific design of user interfaces and selecting suitable interaction patterns depends on the individual application. Therefore, the involvement of the later end-users is of great importance for the development of usable mobile AR solutions.

In [10], Quandt et al. presented a user study to evaluate the subjectively perceived usability and the workload associated with the use of a location-based AR application to support service technicians to conduct maintenance measures on complex heating, air conditioning, and cooling systems on their work location. By taking user feedback into account in the development of the presented AR solution, we improved usability. In addition to the location-based support of service technicians in the work process, which we tested with the system users based on usability and workload evaluation, this article focuses on the learning effects of the users in handling the AR application. Especially concerning the used hand gesture-based user control, further research needs have emerged. In the course of conducting the user study, we observed that users usually learn to use the 3D hand gesture control quickly. After this learning phase, users become more confident in using hand gesture control. Consequently, an optimized design of interaction patterns to fulfill specific work tasks can be concluded. The learning effects that occur when using hand gesture-based interaction will be examined in this paper using the example of the AR application introduced.

Following the related work in chapter 2, we present the case study in chapter 3. In chapter 4, we present and discuss the results of the user studies. The final chapter summarizes the findings of the paper and provides an outlook on further research needs.

2 Related work

In their review, [11] examined the use of AR for industrial application scenarios. The application areas of AR in maintenance deal with the training of employees to perform maintenance tasks, process support for error prevention, maintenance of complex machines and compliance with safety guidelines, and the performance of maintenance activities in hazardous environments. The systematic literature review by [6] provides the state of the art in research on the use of augmented reality to support industrial maintenance activities. The identified state of research includes AR-based assistance systems in various application areas, such as aviation, plant maintenance, or mechanical maintenance. [12] discuss remote support and work process support through virtual information on maintenance objects as core applications for AR in maintenance. In various studies considered for the development of solutions, the focus lies on the tracking procedures used, the mobile AR hardware used, or the interaction between humans and technology. [13] identified great potential to use AR systems for service technicians' training, with the possibility of AR to simulate real work situations. In this context, [14] developed an AR-based learning platform that provides step-by-step instructions for service technicians in the assembly and maintenance of industrial components and plants. An instructor can use the live video image of the trainee to influence the task execution. For location-based learning, [15] have developed an algorithm that identifies real-life learning objects based on the learner's location and provides corresponding learning content. Since many industrial applications require the users' indoor location, the exact localization of the users is a central challenge in order to enable an accurate superimposition of the virtual content. For this purpose, marker-based, SLAM (Simultaneous Localization And Mapping)-based and model-based tracking methods are used in particular. In addition to achieving high accuracy, these methods need to be implemented on mobile AR hardware [16].

Another central challenge in the development of mobile AR-based assistance systems is the interaction with virtual objects in the three-dimensional space. The use of mobile AR hardware alters the requirements for the development of AR user interfaces compared to classical WIMP (Windows, icons, menus, pointers) user interfaces of desktop applications [8]. When using data glasses, hand gesture-based controls are increasingly used. This type of human-technology interaction is particularly suitable for AR applications for direct and, thus, intuitive interaction with virtual objects [17]. The few formal evaluations of hand gesture-based interaction apply user-centered evaluation methods, such as questionnaire-based evaluation of usability and acceptance [17], or by recording and analyzing performance measures from user experiments [18].

To sum up, the challenges for introducing location-based mobile AR applications to support work processes in maintenance are a context-based provision of information, the reliable and accurate recognition of objects, and the use of appropriate kinds of human-computer interaction. In this paper, we focus on the aspect of the experimental testing of a hand gesture-based control system. With the results of the user study, we plan to gain insights for the task-dependent selection of suitable interaction patterns.

3 Case Study

The maintenance of heating, air conditioning, and cooling systems in large infrastructures, such as department stores or airports, places high demands on the service technicians' qualifications. The service technicians' work includes the orientation in the working environment, finding components, documenting measured values, detecting and reporting damages in the context of the maintenance measures carried out. Currently, service technicians conduct maintenance activities with paper-based documents. The service technicians usually carry a maintenance checklist for documentation purposes and a revision plan that contains maintenance components listed in a floor plan of the building. During the execution of maintenance tasks, the search for individual components leads to a considerable loss of time. Due to the often missing documentation of plan changes during component assembly, these search efforts occur. Therefore, updated revision plans can contribute considerably to a more efficient work process. By superimposing the virtual planning basis on the real objects, the trade-specific symbols of the individual components can be displayed and manipulated directly in the field of vision of the service technicians. This way, the service technicians both learn how to work with digitized building data in the work process and contribute to an increased efficiency in the maintenance process through the continuous actualization of the documentation. In this application case, the use of AR glasses offers the advantage of a hands-free usage. Therefore, the technicians' ability carry out maintenance activities is not restricted. Since the service technicians are working indoors, the orientation in the work environment is based on room geometry, derived from the existing building plans.

For this purpose, an importing tool processes the existing revision plans for display on the AR hardware. The importing tool transfers the plans to the mobile hardware according to defined modeling conventions, which, for example, require the arrangement of the room walls on one level of the plan. At the place of maintenance execution, the AR system aligns the virtual revision plan with the real work environment. The user supports the superimposition by setting a starting position that the system matches with the respective revision plan. By moving objects installed at a different location than specified, adding new objects, or deleting objects, the service technicians can directly update the virtual revision plan. After completion of the maintenance task, an export tool prepares the updated revision plans for a subsequent transfer to the order management.

We conducted a user study to evaluate the subjectively perceived usability and the workload associated with the use of the AR application. The results of this study were presented in [10]. By taking user feedback into account in the development of the presented AR solution, we improved its usability. Especially concerning the used hand gesture-based user control, further research needs have emerged. In the course of conducting the user study, we observed that users usually learn to use the 3D hand gesture control quickly. After this learning phase, users generally become more confident in using hand gesture control, and conclusions can be drawn about the design of interaction patterns to fulfill specific work tasks. The learning effects that occur when using hand gesture-based interaction will be examined in this paper using the example of the AR application presented.

4 User tests

At the time of our development of the presented AR application, Microsoft HoloLens™ best met the requirements of the application in the field of maintenance of heating, air conditioning, and cooling systems. The Microsoft HoloLens™ is a semi-transparent Head Mounted Display (HMD) that enables the display of three-dimensional holograms in the user's field of vision based on the reconstruction of the user's real environment [19]. Interaction between the AR hardware and the user bases on the viewing direction (gaze) and hand gestures or voice commands. The user sees a cursor, which he or she controls by head movements, in the center of the field of view that enables the selection of virtual objects by performing a hand gesture, named "air tap." The "air tap" is a hand gesture comparable to the left mouse click and is performed in three steps (see Fig. 1). Through the first two hand movements, the user selects and holds the object (“tap&hold”), and the user can then move it to any position by moving the hand (“tap &move”). By lifting the index finger (gesture 3.), the user rereleases the object. The “air tap” and the resulting hand gestures "tap&hold" and "tap&move" provide the basis for user interaction with virtual objects when using Microsoft Hololens™. At this time, adding individual gestures is not possible when using this hardware.

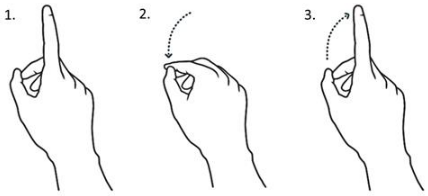


Fig. 1. Execution of the “air tap” hand gesture with the Microsoft HoloLens™ [20]

4.1 Test setup

To investigate the learning effects of using 3D gesture control, we tested five central software functions. After the users set the starting position and the resulting superimposition of the virtual revision plan on the real environment, they used the following manipulation functions: adjust the height of the map display, move an object, duplicate an object, delete an object. A moderator accompanied the test and explained all tasks before the users were carrying them out to ensure comparability. Before performing the tests of the five software functions, all test participants went through a tutorial on how to perform the required hand gestures. After completing the five tasks, all test participants performed these tasks again in the same order. We recorded log files of each user test, including the time required and the number of interactions for the execution of individual tasks as performance measures to evaluate hand gesture-based control's learning effects.

4.2 Task descriptions

The first task, "setting the start position", aims to ensure the accuracy of the superimposition of the virtual building plan with the real working environment. In this step, the user specifies the position in the room and the current viewing direction. To do this, the user performs a "tap&hold" hand gesture after determining his or her position on the room floor plan. A dot appears immediately at the indicated position. The user indicates the viewing direction by pulling out an arrow in the corresponding direction (“tap&move”). By ending the gesture, the user sets the arrow, and herewith confirms the start position or performs the steps to set the start position again (see Fig. 2, left).

The second task, "adjust map height", contains the alignment of the displayed revision plan to the desired height in space, as shown in Fig. 2 on the right. In this case, the task consists of moving the virtual revision plan to the ceiling. To do this, the user performs a "tap&move" hand gesture to grab the virtual map and move it upwards. The user can repeat this hand gesture as often as required to reach the desired height.

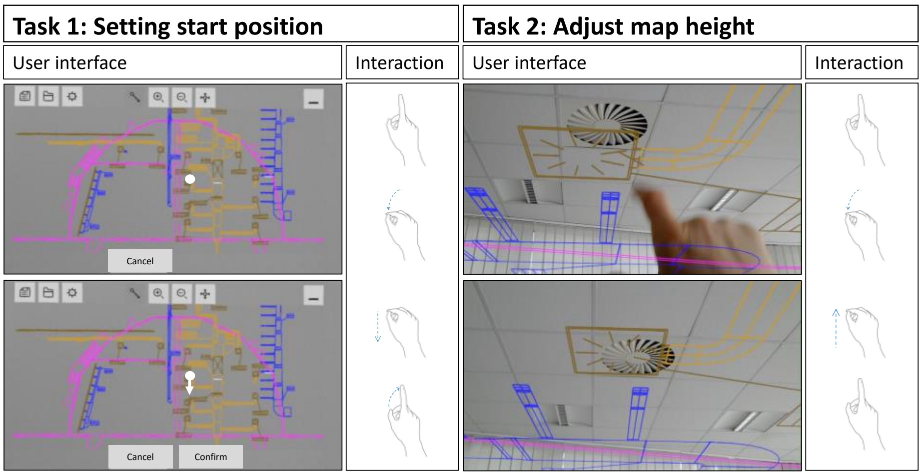


Fig. 2. User interface and interaction pattern for tasks “setting the start position” and “adjust map height” [21]

The following tasks serve to update the revision plan in the working environment. This way, the service technicians learn how to use virtual revision plans and improve the data basis for subsequent maintenance tasks (see Fig. 3).

To complete the third task, "move object", the user moves a selected object of the revision plan from its original position to another position marked with a cross by using a "tap&move" hand gesture. The object can be selected and moved by the user as often as required. If the user moves the object successfully to the target position, the displayed cross disappears, and the user has completed the task.

The fourth task, "duplicate object", is structured as follows: The user marks the object as duplicated by performing a "tap&hold" hand gesture. This way, the user copies the object and moves it to the target position marked with a cross by performing a "tap&move" gesture. In this case, the differences to the task "move object" is not realted to the execution of the hand gestures, but in the representation in the virtual revision plan. After copying the object, the user can select the “move object” mode to adjust the position of the duplicated object as requested.

To fulfill the fifth task, "delete an object," the user marks an object of the virtual revision plan. By executing an "air tap," this object is marked and deleted after the user's confirmation.

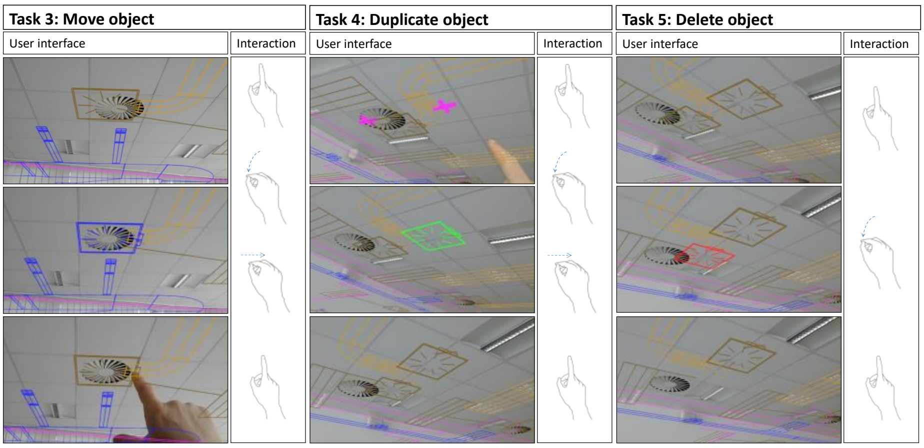


Fig. 3. User interface and interaction pattern for tasks “move object”, “duplicate object”, and “delete object” [21]

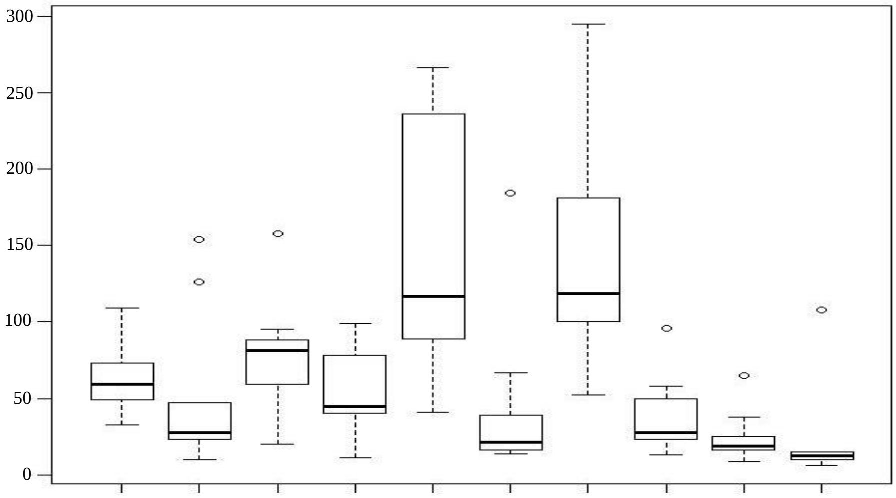
4.3 Composition of the user group

For the user study, we have recruited ten participants, all male (seven students, three academics). The participants were in the age groups 20-25 (three participants), 26-30 (six participants), and 31-35 (one participant). All participants rated their previous experience with computers as high (one participant) or very high (nine participants). The test users rated their previous experience with AR solutions as non-existent (5 participants), first experience (3 participants), or multiple uses (2 participants). In this case, the previous experience of the users had no statistically verifiable effect on the results of the user study, probably due to the sample size. In the maintenance of heating, air conditioning, and cooling systems, the participants estimated their previous experience as non-existent (seven participants), basic knowledge (two participants), and an intermediate level of experience (one participant). Seven participants did not use visual aids; three participants used glasses. Visual aids had no further influence on the test users due to the insertion possibilities of the AR-glasses used. All participants were right-handed.

5 Results and discussion

The results of the user study are shown in Fig. 4 and Fig. 5. Fig. 4 shows the average processing times of the respective tasks for the two test runs. Fig. 5 shows the average user interactions required to complete the task for the two test runs. All test users successfully completed the five tasks in both test runs. This was the prerequisite for us to ensure the comparability of the results.

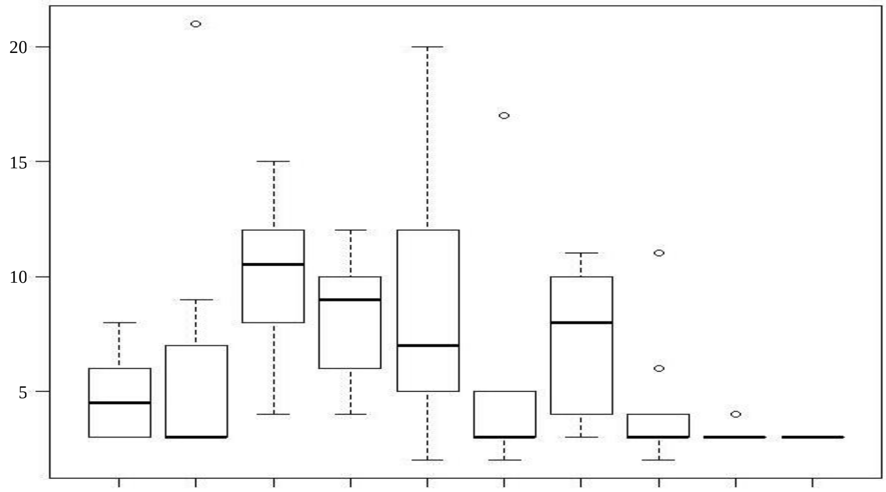
For the first task, "setting the start position", we can determine that the mean processing time to complete the task decreases slightly from the first to the second test run (65 to 48 seconds). However, since the number of required interactions does not considerably reduce, we can observe no learning effects in the use of handheld gesture control in this task. The user's minimum number of three interactions to complete the task is achieved by three users in the first attempt and by seven test users in the second attempt. However, this is in contrast to the very high time and interaction requirements of individual users. The users have to repeat the positioning several times when choosing an inaccurate starting position. Repeated positioning explains the high standard deviation in the processing of this task by the users. From these test results, we conclude that the interaction between users and the developed AR application is not implemented intuitively enough at this point.



	Task 1 - run 1	Task 1 - run 2	Task 2 - run 1	Task 2 - run 2	Task 3 - run 1	Task 3 - run 2	Task 4 - run 1	Task 4 - run 2	Task 5 – run 1	Task 5 – run 2
mean	65,16	48,42	80,19	51,56	142,50	42,42	139,29	37,35	24,36	21,24
median	59,29	27,42	80,98	44,34	116,50	21,51	118,63	27,76	18,89	12,59
standard deviation	23,98	49,63	34,86	27,02	82,09	52,18	71,75	24,84	16,37	30,69

Fig. 4. Box-plot diagrams of processing times for all performed tasks and the two test runs in seconds

Users completed the second task, "adjust map height", faster and with a lower number of interactions compared to the two test runs (see Fig. 4 and 5). Due to the significant reduction of the task processing time by an average of about 35 seconds and a reduction of the average number of interactions by approximately two, we can observe an apparent learning effect in hand gesture control in this task. Depending on the accuracy of the recorded room model, the users had to move the revision plan by about two meters from the starting position to the ceiling. This movement required an average of 10.2 (1st test run) or 8.3 (second test run) interactions. Despite the improved performance measures, we experience the number of "tap&move" hand gestures performed as high for the execution of this task. Therefore, we plan to adjust the movement parameters to allow larger movements of the revision plan along the vertical axis with a gesture's execution. With this adjustment, we expect a further improvement in the performance metrics in the execution of this task.



	Task 1 - run 1	Task 1 - run 2	Task 2 - run 1	Task 2 - run 2	Task 3 - run 1	Task 3 - run 2	Task 4 - run 1	Task 4 - run 2	Task 5 – run 1	Task 5 – run 2
mean	4,90	5,80	10,20	8,30	8,60	4,70	7,20	4,10	3,20	3,00
median	4,50	3,00	10,20	9,00	7,00	3,00	8,00	3,00	3,00	3,00
standard deviation	2,02	5,75	3,12	2,58	5,46	4,45	3,19	2,69	0,42	0

Fig. 5. Box-plot diagrams of number of interactions needed for all performed tasks

In the context of the manipulation tasks "move object", "duplicate object", and "delete an object", we could observe apparent learning effects among the users between the two test runs. Figures 4 and 5 show the corresponding execution times and interaction needs of the users to fulfill the respective tasks. The users were able to reduce the processing time for task 3 by about 70% while reducing the required interactions by about 50%. When looking at the median, this impression manifests as, in the second test run, only one user needed an above-average amount of time to complete the task. In comparing the two test runs, all test users improved both in the time required to complete the task and in the number of required interactions. Due to the analogy of the execution of task 4 compared to task 3, we can observe similar effects in the results. Accordingly, for the execution of task 4, we recorded shorter processing times for all test participants in the second test run. Only two participants needed the same number of interactions in the second run as in the first run; all other participants needed fewer interactions in the second run. The fifth task, "delete an object," does not require any object movement. This task could be performed by almost all participants with the minimum number of interactions, especially in the second test run.

With a critical look at the results of our user study, we are aware that the recorded performance measures do not exclusively reflect the learning effects in dealing with hand gesture-based control. The better understanding of the user's about the tasks and the accuracy of the superposition of the virtual objects with the real working environment influenced the processing time and the number of interactions. Further, the size and composition of the test group can be improved. A higher number of test users and the participation of end-users from the real work environment would have led to more founded and reliable results. Furthermore, for future user studies, the order of tasks could be randomly selected. In this study, the order based on the workflow of the service technicians. In connection with a larger user group, we could have eliminated that the learning effects in using gesture control influenced the processing of individual tasks. The last point to mention is the limited number of hand gestures, which was determined by the selected hardware. The use of other hardware offers different interaction possibilities. Therefore, the results of our study are not necessarily valid across different AR hardware. Nevertheless, we see a clear added value in conducting user studies connected with the development of mobile AR applications for industrial use. This way, essential insights for the design of usable systems can be gained, promoting the acceptance of the developed solutions and thus helping exploit the potentials of AR technology.

6 Conclusion and Outlook

In this paper, we have conducted a user study to investigate the learning effects of using a 3D hand gesture-based control system using the example of an AR application for location-based support of service technicians in the maintenance of heating, air conditioning, and cooling systems. Since there are only few guidelines for designing such human-machine interfaces available, a user-centered evaluation can identify suitable interaction patterns. The five introduced tasks investigated differed in the complexity of the hand gestures to be performed and in the direct relation to 3D virtual objects in space. The user study results confirm that the users learn direct manipulation of virtual objects quickly since the movement of objects with the hands seems intuitive for them. When performing multi-step hand gestures without direct relation to the real environment, we could not detect any learning effects connected with the chosen interaction concept. We believe that multimodal interaction concepts can contribute to a more efficient performance of tasks without an object reference. The testing of such interaction concepts represents a further research requirement for us. Besides, the present study included a limited number of possible hand gestures, which is a result of the hardware selection. By conducting further user studies with a hardware-independent selection of hand gestures, we can transfer the results into general design recommendations for mobile location-based AR assistance systems in the future.

Acknowledgements

The authors would like to thank the German Federal Ministry of Economic Affairs and Energy (BMWi) for their support within the project "Bauen 4.0 – KlimAR – Augmented Reality-based assistance system for the maintenance of complex heating, air conditioning and cooling technology" (grant number 16KN062830).

References

1. Milgram, P. and Kishino, F. A taxonomy of mixed reality visual displays. IEICE Transactions on Information Systems, E77-D(12), (1994).
2. Azuma, R. T. A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments 6(4), 355–385 (1997).
3. Wang, X., Ong, S. K., and Nee, A. Y. C. A comprehensive survey of augmented reality assembly research. Adv. Manuf. 4(1), 1–22 (2016).
4. Kipper, G. and Rampolla, J. Augmented reality. An emerging technologies guide to AR. Syngress, Waltham (2013).
5. Fite-Georgel, P. Is there a reality in Industrial Augmented Reality? In 10th IEEE International Symposium on Mixed and Augmented Reality. IEEE, 201–210 (2011).
6. Palmarini, R., Erkoyuncu, J. A., Roy, R., and Torabmostaedi, H. A systematic review of augmented reality applications in maintenance. Robotics and Computer-Integrated Manufacturing 49, 215–228 (2018).
7. Butz, A. and Krüger, A. (2017. Mensch-Maschine-Interaktion. Walter de Gruyter, Berlin.
8. Dünser, A. and Billinghurst, M. Evaluating Augmented Reality Systems. In: Furht, B. (ed.) Handbook of augmented reality, pp. 289-307. Springer, New York, (2011).
9. Quandt, M., Knoke, B., Gorldt, C., Freitag, M., and Thoben, K.-D. General Requirements for Industrial Augmented Reality Applications. Procedia CIRP. 51st CIRP Conference on Manufacturing Systems, 1130-1135 (2018).
10. Quandt, M., Beinke, T., and Freitag, M. User-Centered Evaluation of an Augmented Reality-based Assistance System for Maintenance. In: Procedia CIRP. 53rd CIRP Conference on Manufacturing Systems, In Print (2020).
11. Bottani, E. and Vignali, G. Augmented reality technology in the manufacturing industry: A review of the last decade. IISE Transactions 51(3), 284–310 (2019).
12. Lamberti, F., Manuri, F., Sanna, A., Paravati, G., Pezzolla, P., and Montuschi, P. Challenges, Opportunities, and Future Trends of Emerging Techniques for Augmented Reality-Based Maintenance. IEEE Trans. Emerg. Topics Comput. 2(4), 411–421 (2015).
13. Oliveira, A. C.M. and Araujo, Regina B., Jardine, Andrew K. S. A Human Centered View on E-Maintenance. Chemical Engineering Transactions 33, 385–390 (2013).
14. Webel, S., Bockholt, U., Engelke, T., Gavish, N., Olbrich, M., and Preusche, C. An augmented reality training platform for assembly and maintenance skills. Robotics and Autonomous Systems 61(4), 398–403 (2013).
15. Tan, Q., Chang William, and Kinshuk. Location-Based Augmented Reality for Mobile Learning: Algorithm, System, and Implementation. The Electronic Journal of e-Learning 13, 138–148 (2015).
16. Bae, H., Walker, M., White, J., Pan, Y., Sun, Y., and Golparvar-Fard, M. Fast and scalable structure-from-motion based localization for high-precision mobile augmented reality systems. mUX J Mob User Exp 5(1), 253 (2016).
17. Shanthakumar, V. A., Peng, C., Hansberger, J., Cao, L., Meacham, S., and Blakely, V. Design and evaluation of a hand gesture recognition approach for real-time interactions. Multimed Tools Appl 79(25-26), 17707–17730 (2020).
18. Ohn-Bar, E. and Trivedi, M. M. Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations. IEEE Trans. Intell. Transport. Syst. 15,6, 2368–2377 (2014).
19. Liu, F. and Seipel, S. Precision study on augmented reality-based visual guidance for facility management tasks. Automation in Construction 90, 79–90 (2018).

20. Microsoft Coporation, https://docs.microsoft.com/de-de/windows/mixed-reality/gestures, last accessed 2020/07/01.
21. Stern, H., Quandt, M., kleine Kamphake, J., Beinke, T., Freitag, M. User Interface Design für Augmented Reality-basierte Assistenzsysteme - Konzept und Anwendungsbeispiele für eine menschorientierte Gestaltung der Nutzer*innenschnittstellen. In: Freitag, M. (ed.): Mensch-Technik-Interaktion in der digitalisierten Arbeitswelt. Schriftenreihe der Wissenschaftlichen Gesellschaft für Arbeits- und Betriebsorganisation. GITO Berlin, In Print (2020).