Asynchronous Manual Work in Mixed Reality Remote Collaboration

Anjela Mayer[1] [0000−0002−5426−3020], Théo Combe[1,2][0000−0003−1209−5580],
Jean-Rémy Chardonnet[2][0000−0002−8926−1359], and Jivka
Ovtcharova[1][0000−0002−7484−5204]

[1] Institute for Information Management in Engineering, Karlsruhe Institute of
Technology (KIT), Karlsruhe, Germany
anjela.mayer@kit.edu
[2] Arts et Metiers Institute of Technology, LISPEN, HESAM Université,
Chalon-sur-Saône, France

Abstract. Research in Collaborative Virtual Environments (CVEs) is becoming more and more significant with increasing accessibility of Virtual Reality (VR) and Augmented Reality (AR) technology, additionally reinforced by the increasing demand for remote collaboration groupware. While the research is focusing on methods for synchronous remote collaboration, asynchronous remote collaboration remains a niche. Nevertheless, future CVEs should support both paradigms of collaborative work, since asynchronous collaboration has as well its benefits, for instance a more flexible time-coordination. In this paper we present a concept of recording and later playback of highly interactive collaborative tasks in Mixed Reality (MR). Furthermore, we apply the concept in an assembly training scenario from the manufacturing industry and test it during pilot user experiments. The pilot study compared two modalities, the first one with a manufacturing manual, and another using our concept and featuring a ghost avatar. First results revealed no significant differences between both modalities in terms of time completion, hand movements, cognitive workload and usability. Some differences were not expected, however, these results and the feedback brought by the participants provide insights to further develop our concept.

Keywords: Asynchronous Remote Collaboration · Collaborative Virtual Environments · Mixed Reality · Asymmetric Collaboration

1   Introduction

Social collaboration is an important component in our daily work where the importance of remote collaboration systems is rapidly growing [9]. The work in companies, institutes and educational facilities increasingly involves stakeholders and interdisciplinary experts from all around the world. To address the effects of globalization, CVEs are required which the participants can virtually join from their remote locations and together conduct collaborative work within the shared virtual environment. In recent years, the covid-19 pandemic has boosted

the transition of working from co-local to remote paradigms, especially during periods of social distancing and lockdowns [1]. As a positive side-effect CVEs help in minimizing travel. Since the stakeholders can simply meet in a shared virtual environment the need for travels to the remote locations of other collaborators decreases which results in a reduction of the carbon footprint. Global collaboration is characterized by spatial distribution of the participants. Additionally, the work force is time-distributed due to different time-zones. The need for CVEs supporting remote asynchronous collaboration increases with a growing global distribution of the collaborating teams. Asynchronous collaboration is required for a successful collaboration [33] and has several unique advantages over synchronous communication, such as: work parallelism, flexible time-coordination, reviewability, and reflection [3, 13, 27, 15]. Nevertheless, CVEs should support synchronous as well as asynchronous collaboration and furthermore allow transition between both [3, 9, 23, 21].

In application areas where the communication of spatial information is important, immersive CVEs are promising platforms to enable effective remote collaboration [15]. Because of advancements in MR technology, collaboration in immersive CVEs has become a research area [9, 3, 18]. De Belen et al. emphasize the need for asynchronous MR CVEs and discuss possible application areas [3].

The literature reviews [3, 9, 18] show an increasing interest in immersive collaboration approaches, in particular the number of publications addressing remote collaboration in MR is growing fast. Nevertheless, while the majority of publications is focusing on synchronous collaboration, the reviewers identified a gap in research of asynchronous collaboration. At the same time they emphasize the benefits of asynchronous CVEs and encourage further research in this area [3, 9, 7].

In this work we propose to follow the suggestions and research application concepts for asynchronous remote collaboration using MR technologies. In particular our main research question is: *How to relive manual work of a non-present collaborator in interactive CVEs?*

As main contribution of this paper we present a concept for the asynchronous record and replay of spatial motions of remote collaborators, in particular their hand motions and interactions with objects within a CVE. Furthermore, we show the application of this concept in an assembly training scenario from the manufacturing area and test it during pilot user experiments.

2   Related Work

This work mainly contributes to the research area of asynchronous CVEs, in particular immersive CVEs which are realized with MR technologies. In asynchronous collaboration scenarios the participants conduct their cooperative work at different times [7]. A key concept in asynchronous collaboration is to create and preserve digital information which can be reconstructed and consumed at another time [3]. In immersive CVEs this often involves the recording of users' actions within the virtual space and a later replay of these actions. To visual-

ize spatial actions of the users within immersive virtual environments they are usually represented by 3D avatars. In this work we use the ghost metaphor as a representation method of non-present user actions.

2.1   Asynchronous collaboration in MR

Asynchronous collaboration in MR is only slowly growing as a research topic, since most of the research in the CVE area is focusing on synchronous collaboration [3, 9, 15, 7]. Ens et al. have reviewed 110 papers about collaboration in MR published between 1995 and 2018. They found that the vast majority of papers (106, or 95%) focus on synchronous collaboration [9]. Their findings are also backed up in the literature review of de Belen et al. [3] where a total of 259 papers between 2013 and 2018 were reviewed. Through this work, we aim at putting our stone to asynchronous collaboration.

Most of the papers about asynchronous collaboration in MR allow the creation and consumption of annotations, like virtual graffiti and photos which are placed at certain locations within the immersive environment and can be viewed and interacted with by other collaborators at another time [17, 24, 19, 6, 31]. Irlitti et al. [16] are researching combination methods for tangible markers and augmented annotations which can be left for the next worker. However, tasks in the engineering domain often involve continuous spatial information which is hard to communicate using static annotations and images.

Tseng et al. [36] present a system which not only preserves respectively correct annotations, but additionally visualizes the position and orientation of the recorded user's head and hands. This provides the minimum of continuous information to perceive the movements of the user's head and hands over time. In the work of Tsang et al. [35] an AR system is developed which can record multimodal streams of annotation data, including viewpoints, voice and gesture information. After a recording is complete, users can save or playback the annotation session.

While the majority of the literature focus on general concepts providing proof-of-concept prototypes, others show how to apply asynchronous collaboration methods to specific domains. A collaboration system for crime scene investigators with remote support from experts is presented by Poelman et al. [29]. Although the main focus is on synchronous collaboration during the investigation, the authors also discuss a record option to support a later review of the investigation research by judges. Marques et al. [21] present a collaboration system which enables remote experts to support on-site technicians with augmented annotations during synchronous as well as asynchronous sessions.

2.2   Spatial capture and replay of body motions

While the creation, preservation and later consumption of information in MR has been considered in existing research, the asynchronous combination of these actions has seldom been considered [15]. V-Mail [14] and MASSIVE-3 [11] are the most relevant approaches where the capture and replay of rich, multi-modal

interactions were applied for asynchronous communication. Chow et al. [7] identified several application domains where this method was implemented for asynchronous collaboration: architectural review [12], creative feedback applications [25, 35], training [38, 4] and tele-communication [6, 28, 30]. In their work Chow et al. present a VR environment enabling asynchronous collaboration in spatial tasks by supporting multi-modal record and replay functionalities and several annotation methods. Other research groups focus on reliving virtual reality experiences and even support the recording and replaying of full body avatars [37, 10]. Although, the literature presents methods for reliving experiences or collaborative planning sessions, further concepts have to be explored suited for the engineering domain, in particular involving tools and manual actions.

Lindlbauer and Wilson [20] propose several time manipulation methods for asynchronous sessions, including pause, loop and replay of a captured 3D environment as well as speed manipulation and jumping back to important moments during meetings. Their methods are useful for applications where users want to make temporal changes. Ogasaware and Shibata [26] developed a prototype CVE system which allows to record user editing to the scene and also to create snapshots of the scene state which can be preserved and operated like non-immersive version control systems such as git.

2.3   Representation of non-present collaborators

To address the loss of physicality in remote work and provide collaborators with awareness about what other collaborators are doing, groupware researchers are exploring user embodiment [9]. Embodiment must represent the functions within the CVE that a collaborator's body and hands would have during his work in the real environment, for instance his hand gestures. In immersive virtual environments users are usually represented by 3D avatars.

The ghost metaphor is a representation style for 3D avatars and was introduced as an intuitive and effective method for training within an immersive environment in the work of Yang and Kim [38]. As they describe, the motion of a trainer is visualized as a *ghost* moving out of the trainee's body in real-time. The trainee spectates the *ghost's* motion from the first-person view and tries to "follow" it as close as possible to imitate the trainer. This kind of interaction is only possible in MR and with additional algorithms the performance of the trainee can be evaluated. In a later work their method is extended with motion-retargeting which converts the recorded trainer data to different body sizes [2]. Their method was tested in experiments for training in fencing, dancing and calligraphy tasks and showed to be as effective as traditional learning methods despite a relatively low presence and problems with MR devices. The motion guidance system of Schönauer et al. [32] expands the *ghost* metaphor to multi-modal guidance feedback adding vibrotactile and pneumatic actuation. Their design space discussion can be used to guide developers of multimodal motion guidance systems. Further research explicitly focuses on arm and hand motion feedback utilizing the *ghost* metaphor [22, 34, 8].

2.4 Contribution

With our work we aim to contribute to the research area of MR asynchronous collaboration. In particular our concept includes a capture of the users actions to preserve their work process and thereby allows to relive this progress again at another time. Unlike past research, we propose a recording of continuous information in which the hands, head and manipulated objects' positions and rotations are kept, without however allowing users to record static information by placing annotations.

Furthermore, we applied our concept to a real use case in the engineering, extending the results in the related work by another application domain. The current study is a first step in the development of our collaborative application, aiming at enabling, for instance, experts to collaborate asynchronously with on-site technicians, or teachers with students.

In the scope of this work we focus on VR, although our concept also works in AR as our first tests with a Microsoft Hololens 2 have shown. Furthermore, our concept can support MR since the recordings of the user could be created in AR and played back in VR and vice versa. This is a first step towards a system which supports the transition between AR and VR as it is encouraged in the literature [9]. The adaptation of our concept to AR and MR will be evaluated in future work.

3 Asynchronous Capture and Replay of Spatial Work

In this work we present a MR concept for asynchronous capture and replay of manual work in remote collaboration. By visualizing former work processes of the collaborators, our concept enables communication through time which is a basic requirement in asynchronous collaboration.
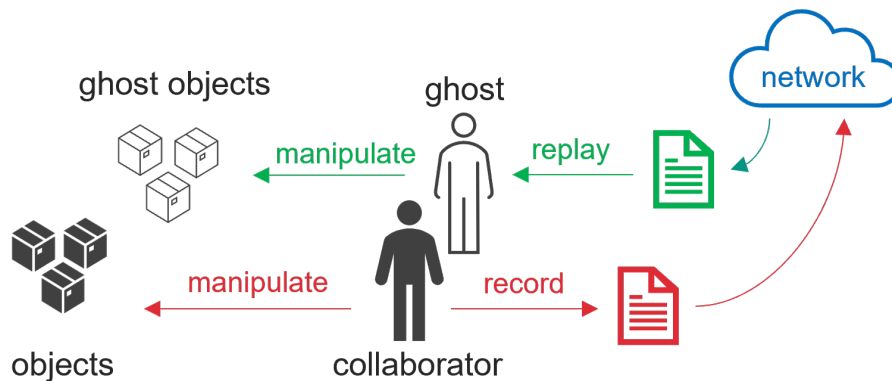

Fig. 1. Overview of the MR recording and replaying process in asynchronous collaboration from the perspective of one collaborator.

As depicted in Figure 1, collaborators can capture their movements and interactions within their environment and send the resulting records to the other participants. Recorded interactions with virtual objects include information about the object's pose and appearance. Received recordings from others can be replayed visualising the non-present collaborators and their interacted objects as *ghosts*. During the MR replay process the collaborators are not restricted in their actions and are able to move and interact freely with their environment while observing the *ghosts*.

3.1   Representation of collaborators and their manual work

To visualize the manual work processes of the collaborators they are represented by 3D avatars. Because we focus on manual work, the representation of the hands and their actions is significant. Nevertheless, a minimal representation of the body is also needed to perceive the presence of the collaborators within the virtual environment.

In VR the hand poses and gestures are tracked via VR controllers. The gestures are visualized by predefined hand stances. In the current implementation three hand stances with the according transition animations are included: *open*, *grab* and *want to grab*, as depicted in Figure 2. If the user moves his hand near a virtual object the virtual hand switches from the *open* to *wants to grab* state visualized by slightly bent fingers. By activating the grab button on the controller, near objects can be picked up with the virtual hand, which will switch to the *grab* state visualized by a closed hand holding the object. In follow-up work more hand gestures can be added with minimal effort.
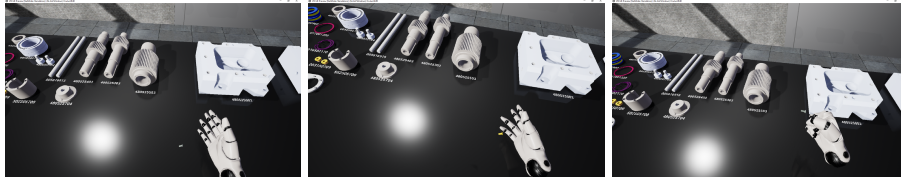

Fig. 2. Left: Open. Middle: Want to grab. Right: Grab stances.

In AR no hand controllers are used. Instead, joint poses of the hands are recognized and mapped onto the virtual 3D hand representation. This method allows a continuous visualization of the hand gestures without discrete states for the hand stances unlike in our VR approach. Furthermore, the user has free hands to interact with the real environment. In contrast to VR, in AR the user is interacting with real objects in his working environment. Thereby only motions of the hands are captured but not of the objects they are manipulating. In future work, image recognition could help to overlay the objects with their according virtual representations and allow recording of their movements in AR as well.

For the visualization of the work process of non-present users, record files are used to reconstruct their 3D avatars and their hand interactions with the virtual objects. To emphasize the absence of the collaborator and improve the visibility of the present environment, the ghost metaphor is utilized. Therefore, a transparent avatar is used for the visualization of the absent collaborator. If the absent user was interacting with virtual objects during the recording, they are represented by transparent duplicates during replay. By creating *ghost* duplicates of the objects, the present user can relive the changes the *ghost* has applied to the virtual environment without actually changing the present environment.

3.2   Record and replay process

In order to reproduce the work of non-present collaborators, the according information has to be collected and saved during their work sessions. The information is frequently collected and saved as a time-frame sequence per user within a record structure as depicted in Figure 3.
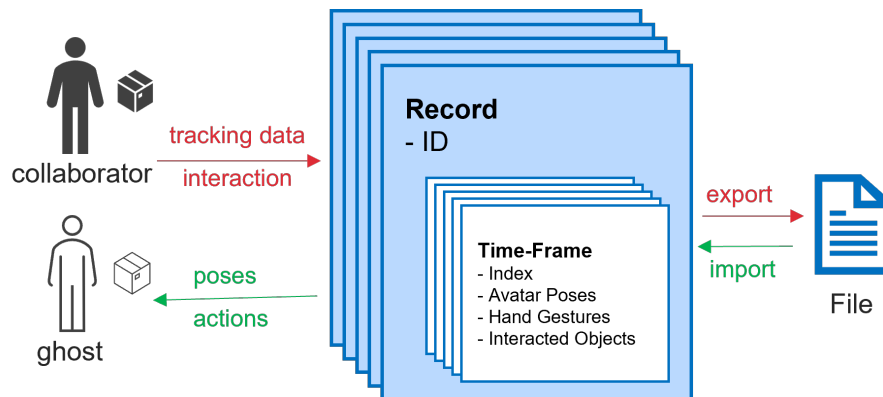

Fig. 3. A record contains sequential information about the collaborator's position, orientation, hand gestures and objects he interacted with. The records can be exported into files to be sent to remote collaborators. Received files can be imported into record data structures which are used to replay the remote collaborator's actions with *ghosts*.

Each time-frame includes information about the avatar poses, hand tracking and interacted objects. The avatar poses consist of the users head and hand positions and orientations within the CVE. Hand tracking information includes the hand states for the gesture representation if recorded in VR and 50 hand joint poses if recorded in AR. For each hand 25 joints are used to construct the gesture of the virtual 3D hand. For the interacted objects the time-frame includes information about the object orientation and position within the CVE as well as information about the appearance of the object, like the 3D model and scale factor.

During replay of a record the contained time-frame sequence is read frame by frame and the according information is applied on the *ghosts*. The avatar poses and hand gestures are applied on the *ghost avatar* representing the absent remote collaborator. If information about interacted objects is available *ghost objects* are created with the according appearance and placed at poses as specified in the record. When the replay is finished all *ghosts* are set to invisible.

For preservation and exchange the records are exported into files, which contain a textual representation of the information. The record files can be sent to the remote collaborators and played within the shared MR environment their local systems are connected to. Therefore, the files should be stored in a format which can be universally used either in VR and AR.

4   Application in Remote Assembly Training

The presented concept for asynchronous remote collaboration was implemented first in a VR assembly training application in manufacturing using the Unreal Engine 4 game engine. Manufacturing involves sequences of manual work steps which consist of spatial information and gestures and are required for a successful assembly completion. The collaborative training scenario takes place between a trainer and trainees where the trainer is showing how to build an assembly within the CVE. The CVE contains all required assembly parts placed on a workbench, which can be picked up and moved with the virtual hands as shown in Figure 4. To support the interaction with the virtual objects a highlighting feature was implemented to visualize the nearest object that can be picked up. For record creation and replay, immersive buttons were implemented which can be pushed with the virtual hands.

In the asynchronous training scenario the trainer is recorded during the assembly process capturing his hand movements and the objects he interacts with. The recording can be activated and stopped by pressing the record button. Once the record is stopped, the data is automatically exported into a text file which can be shared with the trainees.

A trainee only receives the records of the trainer excluding his changes to the CVE with the finished assembly. Being in the initial CVE with the assembly pieces placed on the work bench the trainee can push the play button to view the recording of the trainer and see how to build the assembly as visualized in Figure 4. During replay the ghost of the trainer and the assembly parts he has interacted with are visualized. By observing and imitating the *ghost* the trainee is guided through the assembly process without the simultaneous presence of a trainer. Furthermore, the trainee can record his assembly attempts and send it back to the trainer for review.

5   Pilot Experiment

This section presents a pilot study of our asynchronous collaboration application, considering the VR remote assembly training task detailed above. Thereby we
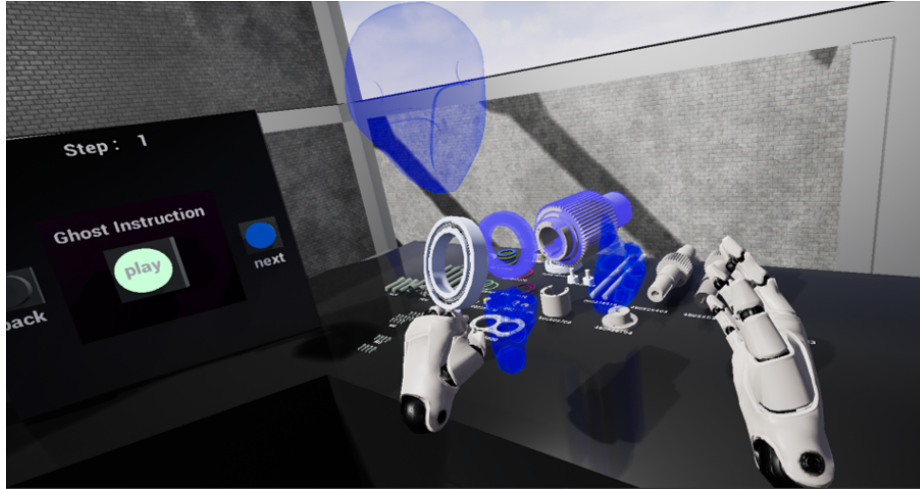
Fig. 4. VR implementation of the asynchronous collaboration concept in an assembly training scenario. The user can interact with the 3D assembly parts using his virtual hands (grey). During replay the *ghosts* (blue) of the trainer and the interacted objects can be viewed.

considered a small sample size and we carried out the experiment using HMD devices, the AR mode being evaluated in a future study. The goal of this pilot study was to get a first insight on our application in a real use case, so that our concept could be enhanced with additional features and fully meet the actual participants' needs.

5.1   Experimental design

10 participants (mean age= 24 ± 3, 1 female) were recruited inside the university, eight of them were familiar with VR devices. The experiment needed approximately thirty minutes to complete.

Upon arrival they were asked to fill in a short demographic questionnaire. Then we explained the purpose of the application, and we gave them the HMD. The HMD used was an *Oculus Quest 2*, which is a lightweight HMD (503g), with a 1832×1920 resolution per eye and a refresh rate of 120 Hz.

Prior starting, we asked them to pick up a part on the workbench and to go through the steps without completing the assembly task, to ensure that they knew and understood how to use the application. Then, we requested them to start the assembly task. The assembly comprised ten steps to complete, and consisted in positioning parts with the others, including ball bearings, screws, circlips, rods and gears. During assembly, we recorded the time needed to complete the whole assembly task, and the hand positions. Once finished, participants were required to leave the application, and to complete three questionnaires: the NASA-TLX (Task load Index), the SUS (System Usability Scale) and a post-

questionnaire about specific points about the application, such as the "ease of use", the "amount of time needed to complete the scenario" and the "satisfaction with the information given to complete the task". Last, participants were free to express themselves about the application, anything that they would improve, or if they were specifically satisfied about the features.

For this pilot study, two modalities were compared: one with the information provided by a *Manual* (M), and one with the information provided by a *Ghost* avatar (G), see Figure 5. In each modality, buttons were displayed on the information panel, participants could push them to move to the next steps. Each participant had to complete one or the other modality, meaning that five of them conducted the M modality and the other five did the G modality.
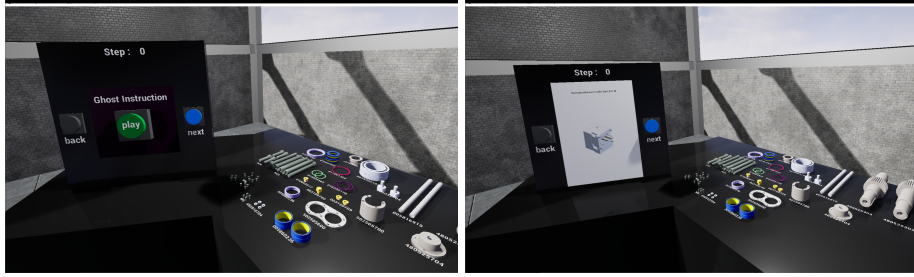

Fig. 5. Left: Ghost modality. Right: Manual modality.

We made the following hypotheses:

H1. The manual modality will induce higher cognitive workload than the ghost modality, because users will have to read the manual first and find the right part on the workbench.

H2. The ghost modality will lead to less hands movements, because users will not have to search for the right part on the table. Moreover participants will make less mistakes in the placement of the parts.

5.2   Results

The experiment led to a between-subject experiment. Each participant completed the application once with one modality, either G or M.

For all the data collected, we performed normality checks. When data were found normal, t-tests were run. On the contrary, when data were found not normal, Mann-Whitney tests were used. The significance threshold was set to .05.

Hand movement: We recorded the positions of the hands at each frame, thus we were able to calculate the amount of movement in meters performed by each participant during the whole application. A t-test on the total distance done by the participant's hands showed no significant differences between

both modalities ($t(8) = 1.869, p = .24$). Nonetheless, the ghost modality led to more movements than the manual modality ($M_G = 350, SD_G = 19.966, M_M = 317.291, SD_M = 40$), see Figure 6 left. This result is unexpected and contradicts hypothesis H2, as we supposed that participants would stand still looking at the ghost and then reproduce the exact movement, while in the *Manual* modality they would have to search for the right part and the right position, which would involve more movements. Furthermore, we made the assumption that participants might imitate the ghost motion by motion, whereas the ghost might perform more movements than necessary, compared to the manual modality. Further investigation is needed to make this point clear.

Completion time: No statistical difference was found in time completion ($t(8) = 1.859, p = .050$). However, since the p-value obtained is quasi-equal to the significance threshold, we could expect clearer results with more participants. Here, participants took in average more time to complete the assembly with the ghost modality, 13.30 minutes ($SD_G = 1.22$min), than with the manual modality, 10.12 minutes ($SD_M = 1.12$min), see Figure 6 right. Time might be higher for the ghost modality, as participants had to wait for the ghost to accomplish each step, which, inherently, doubles the time needed, even though participants successfully accomplished the step at the first attempt. Thus, despite these results, we can assume that the manual modality may take in fact more time than with the ghost, from an absolute perspective.

SUS: A Mann-Whitney test did not reveal any significant difference between both modalities ($U = 6.5, p = .245$). The average scores for both modalities are above 75 ($M_G = 86, SD_G = 3.5, M_M = 78.5, SD_M = 5.624$), which is recognized to show high usability of the application. Only two participants rated under 75 for the manual modality. The scores for each item of the SUS questionnaire is depicted in Figure 7. These results show that our application, whatever the modality, is usable, and none of the participants encountered major challenges when using it, which was one goal of this pilot study. Moreover, since improvements of our concept according to participants' feedback are already planned, the usability score should further increase.

Post-Questionnaire: Three questions were asked: 1) *Overall, I am satisfied with the ease of completing the tasks in this scenario*, 2) *Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario*, 3) *Overall, I am satisfied with the information media when completing the tasks*. Participants had to answer on a Likert scale ranging from 0-Strongly Disagree to 5-Strongly Agree. Although both modalities provided satisfaction to participants ($M_G = 4.4, SD_G = .163, M_M = 4.6, SD_M = 0.163$), no significant differences were found between each modality ($U = 8.5, p = .403$).

NASA-TLX: We asked participants to complete the NASA-TLX upon completion, no significance difference could be highlighted ($t(10) = 1.814, p = .121$). If we look closely at each item of the NASA-TLX, we can observe that except for *Performance* and *Frustration*, each item has been rated higher for modality M, the higher difference being for *Mental Demand*, as depicted in Figure 8. Although not significant, these observations tend to go in the same direction as
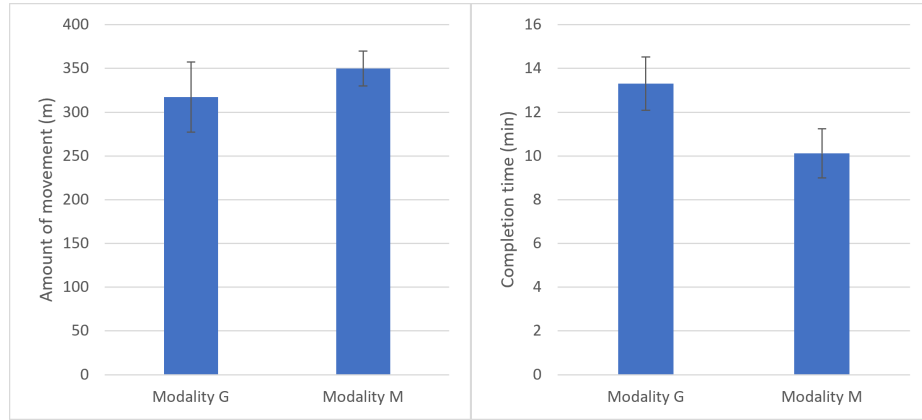
Fig. 6. Left: Hand movements. Right: Completion time.

hypothesis H1. Further investigation is however needed, as we took participants not coming from industry, for whom reading a manufacturing manual may be more cognitive demanding than for industrial operators.

Participant feedback and observation: During the experiment, we observed the participants, and after it we ask them to provide feedback. First of all, participants performing the manual modality tended to consider gravity and leave the parts directly on the workbench, even though it meant that they had to turn themselves to access a specific viewpoint, or they kept the part in their hand during the whole process. Moreover, four of them used only one hand. On the contrary, users in the ghost modality used both hands and were more likely to ignore gravity and leave parts at their eyes' height. We may suppose that they took inspiration from what the ghost avatar did, which was to leave objects in the air and use his both hands. Participants gave advice on how to improve the application, such as having a time progression bar when the ghost performs a step, to know how much time left there would be before the ghost would stop, or adding physical feedback when a part is grabbed (currently the parts are highlighted in yellow when grabbed) or when they accurately place the part they are holding. Some steps require unnatural placement movements, such as placing screws which normally requires specific tools. For such steps, tools, such as screw drivers, could be added to enhance realism.

5.3   Limitations

Several limitations might have influenced the results obtained. First, except one participant, all participants were from our research institute, which activity focuses on MR, thus, they all had consequent experience with VR tools, which is not representative of final users. Furthermore, our results could be influenced by the novelty effect, which must be minimized in VR training applications [5]. Our post-questionnaire showed a slightly greater satisfaction of inexperienced par-
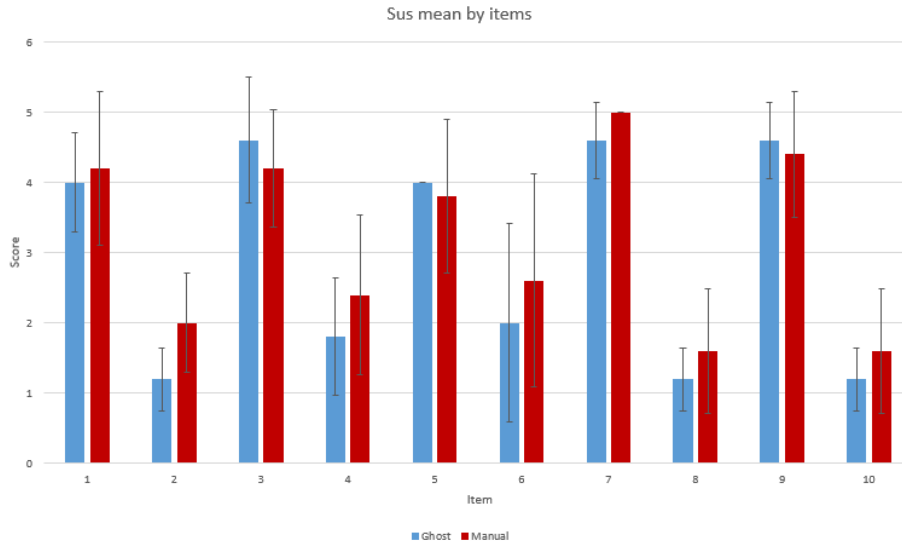
Fig. 7. SUS items results.

ticipants compared to participants who had already spent more than 20 hours in VR. Additionally, inexperienced subjects were more satisfied with the ghost modality compared to the manual. On the contrary, for experienced participants as well as in the overall results, a greater satisfaction was observed with the manual modality. Nevertheless, to be statistically more representative, a larger population sample should be considered. At last, an additional possible source of error is the current state of the application, as in fact, some improvements, such as the ones described above, may lead to significant changes in performance.

6   Conclusion

In this work we presented a concept for asynchronous remote collaboration based on the recording and replaying of body motions, in particular hand motions of the collaborators during manual work, in MR. The concept was applied in a remote assembly training scenario in VR and tested in a pilot experiment featuring two modalities: one with a manual displayed next to the workbench, and one with our ghost recording. This experiment has revealed no significant differences between both modalities in terms of hands movements, completion time, cognitive workload and usability. However, during the experiment, we observed clear differences in behaviors, which may impact past results with tasks requiring more steps. Moreover, we had only five participants for each modality, which may have impacted the results. We believe that more significant differences could be observed in a future experiment involving more steps to perform and integrating
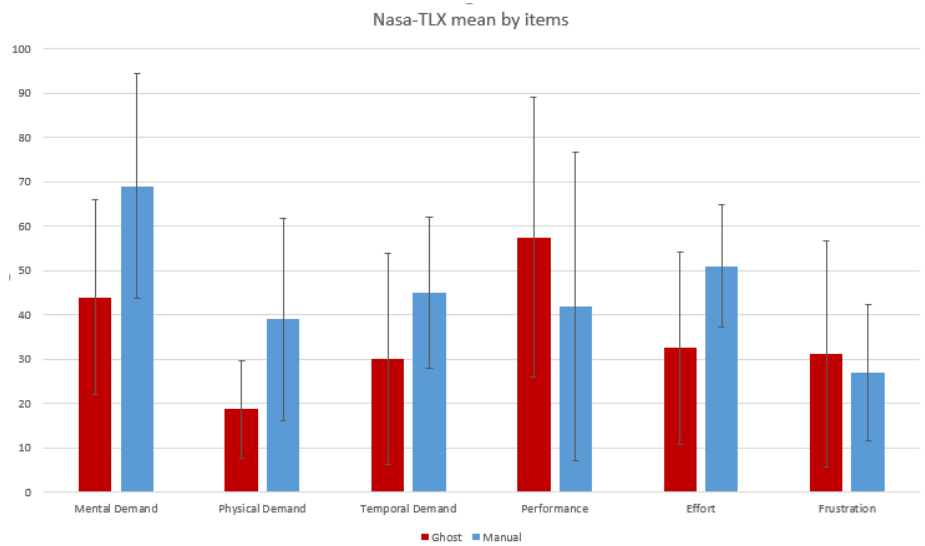
Fig. 8. NASA-TLX items results.

in the application the features proposed by the participants. Nonetheless, this pilot experiment provides insights to further develop the proposed concept.

For the evaluation of our concept a larger user study is planned as well as further experiments including AR. In the future work we aim to build a system for asynchronous MR collaboration including the presented record and replay of the collaborators as well as a synchronization of the shared CVE between all participants to exchange manipulations of the virtual environment along with the records. Additionally, we will work on multi-modal recordings, including verbal communication, pursue new concepts for time navigation through the recordings and implement our concepts in further application domains.

References

1. Almeida, F., Duarte Santos, J., Augusto Monteiro, J.: The challenges and opportunities in the digitalization of companies in a post-covid-19 world. IEEE Engineering Management Review **48**(3), 97–103 (2020). https://doi.org/10.1109/EMR.2020.3013206
2. Baek, S., Lee, S., Kim, G.J.: Motion retargeting and evaluation for VR-based training of free motions. Visual Computer **19**(4), 222–242 (2003). https://doi.org/10.1007/s00371-003-0194-2
3. de Belen, R.A., Nguyen, H., Filonik, D., Favero, D., Bednarz, T.: A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. AIMS Electronics and Electrical Engineering **3**, 181–223 (2019). https://doi.org/10.3934/ElectrEng.2019.2.181

4. Chan, J.C., Leung, H., Tang, J.K., Komura, T.: A virtual reality dance training system using motion capture technology. IEEE Transactions on Learning Technologies **4**(2), 187–195 (2011). https://doi.org/10.1109/TLT.2010.27

5. Checa, D., Miguel-Alonso, I., Bustillo, A.: Immersive virtual-reality computer-assembly serious game to enhance autonomous learning. Virtual Reality (2021). https://doi.org/10.1007/s10055-021-00607-1

6. Chen, H., Lee, A.S., Swift, M., Tang, J.C.: 3d collaboration method over hololens™ and skype™ end points. In: Proceedings of the 3rd International Workshop on Immersive Media Experiences. p. 27–30. ImmersiveME '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2814347.2814350

7. Chow, K., Coyiuto, C., Nguyen, C., Yoon, D.: Challenges and design considerations for multimodal asynchronous collaboration in VR. Proceedings of the ACM on Human-Computer Interaction **3**(CSCW) (2019). https://doi.org/10.1145/3359142

8. Dürr, M., Kn, M.D., Weber, R., Pfeil, U., Reiterer, H.: EGuide: Investigating different Visual Appearances and Guidance Techniques for Egocentric Guidance Visualizations (2020). https://doi.org/10.1145/3374920.3374945

9. Ens, B., Lanir, J., Tang, A., Bateman, S., Lee, G., Piumsomboon, T., Billinghurst, M.: Revisiting collaboration through mixed reality: The evolution of groupware. International Journal of Human-Computer Studies **131** (2019). https://doi.org/10.1016/j.ijhcs.2019.05.011

10. Frécon, E., Nöu, A.A.: Building distributed virtual environments to support collaborative work. Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST pp. 105–114 (2019). https://doi.org/10.1145/293701.293715

11. Greenhalgh, C., Flintham, M., Purbrick, J., Benford, S.: Applications of temporal links: recording and replaying virtual environments. In: Proceedings IEEE Virtual Reality 2002. pp. 101–108 (2002). https://doi.org/10.1109/VR.2002.996512

12. Guerreiro, J., Medeiros, D., Mendes, D., Sousa, M., Jorge, J., Raposo, A., Santos, I.: Beyond Post-It: Structured Multimedia Annotations for Collaborative VEs. In: Nojima, T., Reiners, D., Staadt, O. (eds.) ICAT-EGVE 2014 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments. The Eurographics Association (2014). https://doi.org/10.2312/ve.20141365

13. Hollan, J., Stornetta, S.: Beyond being there. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. p. 119–125. CHI '92, Association for Computing Machinery, New York, NY, USA (1992). https://doi.org/10.1145/142750.142769

14. Imai, T., Johnson, A., Leigh, J., Pape, D., DeFanti, T.: Supporting transoceanic collaborations in virtual environment. pp. 1059 – 1062 vol.2 (1999). https://doi.org/10.1109/APCC.1999.820446

15. Irlitti, A., Smith, R.T., Itzstein, S.V., Billinghurst, M., Thomas, B.H.: Challenges for Asynchronous Collaboration in Augmented Reality. Adjunct Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2016 pp. 31–35 (2017). https://doi.org/10.1109/ISMAR-Adjunct.2016.0032

16. Irlitti, A., Von Itzstein, S., Alem, L., Thomas, B.: Tangible interaction techniques to support asynchronous collaboration. In: 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 1–6 (2013). https://doi.org/10.1109/ISMAR.2013.6671840

17. Kasahara, S., Heun, V., Lee, A.S., Ishii, H.: Second surface: Multi-user spatial collaboration system based on augmented reality. In: SIGGRAPH Asia 2012 Emerging Technologies. p. 1–4. SA '12, Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2407707.2407727

18. Ladwig, P., Geiger, C.: A literature review on collaboration in mixed reality. In: Auer, M.E., Langmann, R. (eds.) Smart Industry & Smart Education. pp. 591–600. Springer International Publishing, Cham (2019)

19. Langlotz, T., Grubert, J., Grasset, R.: Augmented reality browsers: Essential products or only gadgets? Commun. ACM **56**(11), 34–36 (2013). https://doi.org/10.1145/2527190

20. Lindlbauer, D., Wilson, A.D.: Remixed Reality: Manipulating Space and Time in Augmented Reality, p. 1–13. Association for Computing Machinery, New York, NY, USA (2018)

21. Marques, B., Silva, S., Rocha, A., Dias, P., Santos, B.S.: Remote asynchronous collaboration in maintenance scenarios using augmented reality and annotations. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). pp. 567–568 (2021). https://doi.org/10.1109/VRW52623.2021.00166

22. Matsas, E., Vosniakos, G.C., Batras, D.: Modelling simple human-robot collaborative manufacturing tasks in interactive virtual environments. In: Proceedings of the 2016 Virtual Reality International Conference. VRIC '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2927929.2927948

23. Moloney, J., Amor, R.: Stringcve: Advances in a game engine-based collaborative virtual environment for architectural design. In: In Proceedings of CONVR 2003 Conference on Construction Applications of Virtual Reality. pp. 24–26 (2003)

24. Nassani, A., Bai, H., Lee, G., Billinghurst, M.: Tag it! ar annotation using wearable sensors. In: SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications. SA '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2818427.2818438

25. Nguyen, C., DiVerdi, S., Hertzmann, A., Liu, F.: Collavr: Collaborative in-headset review for vr video. In: Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. p. 267–277. UIST '17, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3126594.3126659

26. Ogasawara, H., Shibata, Y.: Asynchronous collaborative support system by revision tree method. Proceedings - International Conference on Advanced Information Networking and Applications, AINA pp. 564–569 (2009). https://doi.org/10.1109/WAINA.2009.180

27. Olson, G.M., Olson, J.S.: Distance matters. Human–Computer Interaction **15**(2-3), 139–178 (2000). https://doi.org/10.1207/S15327051HCI1523_4

28. Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., Kim, D., Davidson, P.L., Khamis, S., Dou, M., Tankovich, V., Loop, C., Cai, Q., Chou, P.A., Mennicken, S., Valentin, J., Pradeep, V., Wang, S., Kang, S.B., Kohli, P., Lutchyn, Y., Keskin, C., Izadi, S.: Holoportation: Virtual 3d teleportation in real-time. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology. p. 741–754. UIST '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2984511.2984517

29. Poelman, R., Akman, O., Lukosch, S., Jonker, P.: As if being there: Mediated reality for crime scene investigation. In: Proceedings of the ACM

2012 Conference on Computer Supported Cooperative Work. p. 1267–1276. CSCW '12, Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2145204.2145394

30. Regenbrecht, H., Meng, K., Reepen, A., Beck, S., Langlotz, T.: Mixed voxel reality: Presence and embodiment in low fidelity, visually coherent, mixed reality environments. pp. 90–99 (2017). https://doi.org/10.1109/ISMAR.2017.26

31. Reisner-Kollmann, I., Aschauer, A.: Design and implementation of asynchronous remote support. CEUR Workshop Proceedings **2618**, 9–11 (2020)

32. Schönauer, C., Fukushi, K., Olwal, A., Kaufmann, H., Raskar, R.: Multimodal motion guidance: Techniques for adaptive and dynamic feedback. ICMI'12 - Proceedings of the ACM International Conference on Multimodal Interaction pp. 133–140 (2012). https://doi.org/10.1145/2388676.2388706

33. Tam, J., Greenberg, S.: A framework for asynchronous change awareness in collaborative documents and workspaces. International Journal of Human-Computer Studies **64**(7), 583–598 (2006). https://doi.org/10.1016/j.ijhcs.2006.02.004

34. Tecchia, F., Alem, L., Huang, W.: 3d helping hands: A gesture based mr system for remote collaboration. In: Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry. p. 323–328. VRCAI '12, Association for Computing Machinery, New York, NY, USA (2012). https://doi.org/10.1145/2407516.2407590

35. Tsang, M., Fitzmzurice, G.W., Kurtenbach, G., Khan, A., Buxton, B.: Boom Chameleon: Simultaneous capture of 3D viewpoint, voice and gesture annotations on a spatially-aware display (2003)

36. Tseng, P.Y., Haraldsson, H., Belongie, S.: Annotate all ! A Perspective Preserved Asynchronous Annotation System for Collaborative Augmented Reality pp. 1–3 (2019)

37. Wang, C.Y., Sakashita, M., Ehsan, U., Li, J., Won, A.S.: Again, Together: Socially Reliving Virtual Reality Experiences When Separated, p. 1–12. Association for Computing Machinery, New York, NY, USA (2020)

38. Yang, U., Kim, G.J.: Implementation and evaluation of "just follow me": An immersive, vr-based, motion-training system. Presence: Teleoper. Virtual Environ. **11**(3), 304–323 (2002). https://doi.org/10.1162/105474602317473240