

# A Survey on Synchronous Augmented, Virtual and Mixed Reality Remote Collaboration Systems

ALEXANDER SCHÄFER, TU Kaiserslautern

GERD REIS, German Research Center for Artificial Intelligence

DIDIER STRICKER, German Research Center for Artificial Intelligence, TU Kaiserslautern

Remote collaboration systems have become increasingly important in today's society, especially during times where physical distancing is advised. Industry, research and individuals face the challenging task of collaborating and networking over long distances. While video and teleconferencing are already widespread, collaboration systems in augmented, virtual, and mixed reality are still a niche technology. We provide an overview of recent developments of synchronous remote collaboration systems and create a taxonomy by dividing them into three main components that form such systems: *Environment*, *Avatars*, and *Interaction*. A thorough overview of existing systems is given, categorising their main contributions in order to help researchers working in different fields by providing concise information about specific topics such as avatars, virtual environment, visualisation styles and interaction. The focus of this work is clearly on synchronised collaboration from a distance. A total of 87 unique systems for remote collaboration are discussed, including more than 100 publications and 25 commercial systems.

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Human-centered computing** → **Computer supported cooperative work**; **Mixed / augmented reality**; **Virtual reality**; *Collaborative interaction*.

Additional Key Words and Phrases: virtual reality, augmented reality, mixed reality, collaboration, remote assistance, distant cooperation, literature review

## ACM Reference Format:

Alexander Schäfer, Gerd Reis, and Didier Stricker. 2022. A Survey on Synchronous Augmented, Virtual and Mixed Reality Remote Collaboration Systems. *ACM Comput. Surv.* 1, 1, Article 1 (January 2022), 28 pages. <https://doi.org/10.1145/3533376>

## 1 INTRODUCTION

Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR) technologies are becoming more mature and open new ways for remote collaboration. Video and teleconferencing systems are already in extensive use in today's society, enabling a focus on more novel alternatives which utilize virtual, augmented and mixed reality technologies. Systems for remote collaboration in AR, VR and MR are developing slowly but steadily, yet they are not well established. An early example for remote collaboration using a virtual environment is Second Life. Almost two decades ago, the Second Life project [86, 88] allowed universities, corporations, cities, artists and individuals to create and share virtual spaces with many people [48].

---

Authors' addresses: Alexander Schäfer, TU Kaiserslautern, [alexander.schaefer@dfki.de](mailto:alexander.schaefer@dfki.de); Gerd Reis, German Research Center for Artificial Intelligence, [gerd.reis@dfki.de](mailto:gerd.reis@dfki.de); Didier Stricker, German Research Center for Artificial Intelligence, TU Kaiserslautern, [didier.stricker@dfki.de](mailto:didier.stricker@dfki.de).

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0360-0300/2022/1-ART1 \$15.00  
<https://doi.org/10.1145/3533376>

Such systems consist of many different facets and requires expertise in various fields such as 3D modeling, animation, development of interactive systems, creation of avatars and dynamic content, multi-user communication and more. In addition, mixed reality systems are often implemented using a combination of AR and VR hardware, which requires a certain expertise in both technologies.

In this paper we give a thorough summary and discussion of synchronous remote collaboration systems which utilize AR/VR/MR technology. The importance of such systems was emphasized during the COVID-19 outbreak in December 2019. People all over the world were put into quarantine, cities and local communities forbid traveling and even stepping outside. During this time, the scientific community was forced to find novel ways to network and communicate. Most scientific conferences were either cancelled or held completely virtual, some even using immersive 3D experiences utilizing VR HMD's. The IEEE VR 2020 conference as an example, used virtual rooms where conference participants could join and interact with each other. Paper and poster presentations were done within a virtual environment which was streamed online for a broader audience. Although video and teleconferencing systems in particular experienced a significantly increased use as a result of this global crisis, the AR, VR and MR community received a major awareness push as well.

A well sophisticated AR/VR/MR system could help to reduce travel costs, office space, time, and carbon emissions by creating shared immersive spaces with believable person embodiment and interaction. To compete with each other in this crisis, many companies and researchers have recently invested in creating novel systems, which makes a recent review of existing systems and research even more interesting.

We identify and classify the individual parts which are necessary for a remote collaboration system and provide an overview of existing systems for research and professional work. Remote collaboration systems utilizing AR/VR/MR technology are used in many different fields such as human computer interaction, computer graphics, medicine, training, cognitive sciences and many more.

We define three components that each remote collaboration system needs to implement, namely *Environment*, *Avatars* and *Interaction*. A detailed explanation about this taxonomy is described in section 2.4. By providing condensed information on certain key topics, we want to help researchers assess the state of the art in a particular subject. As an example, a researcher focused on novel environments for remote collaboration can use Table 1 which lists the discussed work with respect to their technology, use case, visualisation style and the stimulated sensory inputs. The representation of other users, in regard to their visualisation and animation style is shown in Table 2. A researcher focused on interaction in multi-user collaborative environments will be interested in inspecting Table 3 which categorises important works in regard to interaction types which were found during the survey.

## 2 RELATED WORK AND SURVEY PROCEDURE

Related work was conducted by Phon et al. [71] which reviewed the state of the art in collaborative AR systems with focus on education in 2014. This work has a clear focus on collaborative learning in AR and does not differentiate between remote and local collaboration experiences. Another survey was done by Wang et al [123] with focus on AR and MR based collaborative design in manufacturing. Ens et al. [17] review published work in collaboration through mixed reality up to the year 2018. Although the focus of mentioned work is not remote collaboration explicitly, the authors differentiate between remote and physically co-located systems. Another Survey was conducted by de Belen et al. [10] in which the authors provide a systematic review of collaborative mixed reality technologies. De Pace et al. [11] provides a systematic review for published work in the context of collaborative industrial robots using AR technology. The authors assess relevant

papers and list whether these have been evaluated in a user-centred way and rate the different papers according to their quality criterias. Furthermore, the authors mention that relevant papers should be properly evaluated from a user-centred perspective to truly assess their effectiveness.

Compared to the mentioned surveys, this work has a narrow focus on *synchronous* and *remote* collaboration systems which use either AR, VR or MR technology. This survey categorises relevant papers in a unique way and aims to provide a novel perspective on the current state of the art. Section describes 2.1 which work is considered relevant and section 2.2 the survey procedure.

## 2.1 Inclusion and Exclusion of Systems

In our work we provide a concise focus on *synchronous remote collaboration systems* and we categorise the results to assist scientists in different fields to cover their specific research interest. We emphasize the term **remote**, which means that physical co-location of users is not required and the term **synchronous** which allows users to collaborate in real-time. The focus is on virtual, augmented and mixed reality systems. Traditional video and teleconferencing systems are omitted. We define a remote collaboration system as a way of communicating, interacting and sharing a space beyond the boundaries of physical space exclusively through technological channels with distributed users. We include work which uses or implements a combination of AR/VR/MR technology and synchronous remote collaboration. Systems which allow multiple users in a system but require users to be in physical co-location are excluded. Exceptions are systems which were used in a physical co-location scenario but could easily be extended for remote collaboration purposes. Asynchronous systems that do not allow real-time communication between users are also excluded.

## 2.2 Methodology

The survey was conducted through an iterative process by integrating the most relevant papers first, identifying specific similarities and differences with subsequent categorisation. By incrementally adding new relevant research work we evolved the categorisation process and therefore separated the relevant work into three main contribution categories: *Environment*, *Interaction* and *Avatars*. With this approach it is possible to filter the relevant papers (over 2.000 unique papers) by applying the constraints mentioned in section 2.1 and then fitting the remaining papers into categories of the specific research interest for researchers.

Our extensive search was performed by using search queries in different data sources. The used data sources include Scopus (<https://www.scopus.com>), Google Scholar (<https://scholar.google.com/>), ACM Digital Library (<https://dl.acm.org/>), IEEE Xplore (<https://ieeexplore.ieee.org/>), Springer Link (<https://link.springer.com/>), PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) and Microsoft Academic (<https://academic.microsoft.com/>). Additionally we looked through the proceedings of multiple leading AR and VR conferences such as IEEE VR, ISMAR and EuroVR/EuroXR with focus on collaboration related topics. The search was performed by concatenating AR/VR/MR with keywords such as *remote*, *collaboration*, *social* and more (as shown in Figure 1). The tables in the following chapters summarise works from the same author if it is a continuation or extension of the previous work. Furthermore, the focus is on the general implementation of the proposed systems rather than on their specific research questions.

## 2.3 Definition AR, VR and MR systems

In this section we briefly introduce how we perceive AR, VR and MR systems. In general, AR is achieved with two approaches: *Video-Seethrough* and *Optical-Seethrough*. In both cases, the real world is augmented to the user. In Video-Seethrough the world is captured with a camera and virtual objects are placed onto the captured images. In Optical-Seethrough systems, users perceive

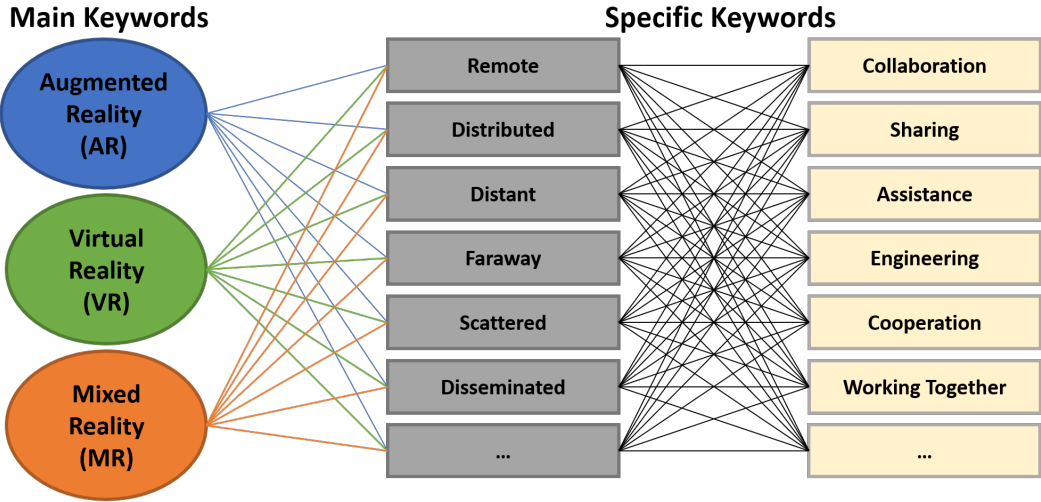


Fig. 1. The used search methodology.

the outside world with their own eyes through a transparent projection surface which displays the AR content. Regarding VR, most literature couples the term VR with a Head Mounted Display (HMD) which is placed on the head of a user. In this survey we also consider systems without HMD's as VR, independent of the specific display device, as long as it is possible to immerse users into a virtual 3D environment. While AR and VR systems are often quite clear in their separation, the distinction among MR systems often leads to confusion. For this survey, we use Milgram et al.'s [62] definition of the Reality-Virtuality continuum which describes mixed reality as the area where both, the real and virtual world are mixed (see Figure 2).

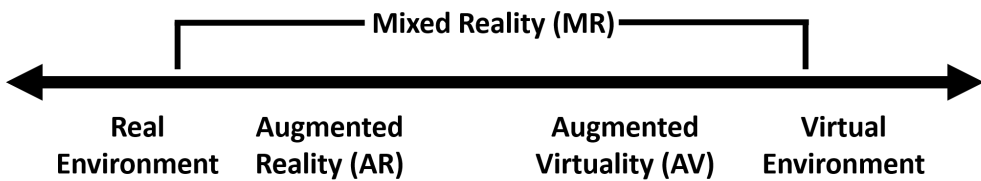


Fig. 2. The Reality-Virtuality continuum according to Milgram et al. [62].

More precisely, a system is referred to as MR in this survey if at least one of the following points applies:

- (1) There is a mix between AR and VR hardware (this also includes projector based systems)
- (2) Real world objects are used for interaction with either AR or VR hardware

Although the recently coined term Extended Reality XR is very popular, we do not use it in this context because we explicitly exclude "pure reality" and the remaining parts are covered by the terms AR, VR and MR.

## 2.4 The three pillars for remote collaboration systems: A taxonomy of remote collaboration systems

We create a taxonomy and categorise the relevant work in a logical manner. Many systems have different aspects of novelty which cannot be described by assigning them to a specific category. As an example, one system might excel in the novelty of avatars while another introduces a new kind of interaction technique for remote collaboration. One goal of this survey is to help researchers from a wide range of fields who are interested in the area of remote collaboration systems which utilize AR, VR and MR technology. To illustrate this: A researcher who is interested in the topic remote collaboration using AR/VR/MR might ask "How are users represented in virtual environments?", "What kind of interaction is possible in a shared virtual space?" or "Are there collaboration systems which enable shared gaze awareness?". With our survey we want to provide condensed information to different research questions such as virtual representation of users, different types of interaction and the virtual environment. To achieve this, we elaborated a concept that enables the possibility to view each of these systems from different viewpoints: *Environment*, *Avatars* and *Interaction* (see Figure 3) which we call **the three pillars of remote collaboration systems**. In the next chapters, we explain each of these components in more detail and present important and highly cited publications in each category. Additionally we provide tables for each category to allow quick access to the desired work: Table 1 is summarizing remote collaboration systems with focus on virtual environment, Table 2 focuses on user representation and Table 3 identifies and categorises different interaction possibilities.

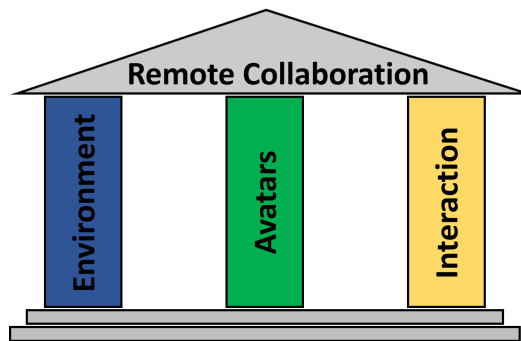


Fig. 3. The three pillars for remote collaboration systems.

## 3 COMPONENTS OF REMOTE COLLABORATION SYSTEMS

### 3.1 Environment

*Virtual environment* refers to a simulated environment that stimulates the sensory impressions of a user. One of the first virtual environments was *Sensorama*, created by Morton Heilig in the early sixties [35]. It featured a simulated motorcycle ride with 3D visuals, stereo sound, olfactory cues (aromas) and tactile cues (seat vibration and wind from fans). In recent literature, most virtual environments are not as comprehensive and complete as the prototype created by Heilig, but rather focus on specific areas that are mostly visual or acoustic stimuli. Exceptions are augmented reality systems which utilize markers, where tangible interfaces with haptic feedback are still popular. As an example, Wang et al. used tangible interfaces such as a regular table [125] or tabletop [127, 128]. Other marker based systems used turntables, such as Shen et al. [95, 96] or additional interaction tools such as a pen in [97]. With increasing maturity of AR technology, marker based systems became obsolete and such systems are not further developed.

In case of VR, there is usually a 3D modeled scene which is rendered, while in AR, the virtual environment refers to the augmented virtual objects superimposed onto the real world. Some AR systems do not include any 3D object rendering but use shared annotations and virtual pointers instead [25–27, 30–32, 55, 89, 90]. Sense of presence, often called telepresence, is highly affected by the quality and consistency of the virtual environment [132]. In early work, studies suggested that the overall sense of presence is increased by adding tactile and auditory cues [13]. The more sensory impressions are added, the greater the feeling of presence according to the studies of Dinh et al. [13]. In VR as example, telepresence is not only achieved by highly realistic 3D environments but also with consistency i.e. avatars should blend in with the environment and interaction methods should be adequate [132]. The work of Yoon et al. [132] compares different types of avatars with different styles. The authors’ findings include that a virtual environment in cartoon style should also use avatars in cartoon style to achieve a higher sense of presence for the users. In Table 1 we present research works, summarised by their respective virtual environment properties and ordered according to their respective technology (AR/VR/MR). Furthermore, we categorised remote collaboration systems in three main use cases: *Meeting*, *Design* and *Remote Expert* since they were most popular and consistent throughout the literature. The category *Meeting* can also be seen as a means of *sharing a workspace with other users*. Some systems which are used for training or socializing fit also in this category. Note that the category *Event* has been added for VR-based systems, as there were three systems that could not otherwise be meaningfully categorised. These systems are used solely for event purposes. Bigscreen [42] focuses on virtual cinemas, allowing people to buy tickets and then watch movies together in a collaborative virtual environment. Sansar [8] and Wave [129] focus on virtual live events such as concerts. To the best of our knowledge, there is no existing system purely designed for events based on AR/MR technology.

System / Authors	Techn.	Use Case	Visualisation Style	Sensory Inputs
Orts et al. [68]	AR	Meeting	Realistic	Audio, Visual
Regenbrecht et al. [85]	AR	Meeting	Realistic	Audio, Visual
Shen et al. [95–98]	AR	Design	Cartoon	Audio, Visual, Tactile
Poppe et al. [77, 78]	AR	Design	Cartoon	Audio, Visual
Sodhi et al. [102]	AR	Remote Expert	Cartoon	Visual
Gurevich et al. [31, 32]	AR	Remote Expert	Annotations	Audio, Visual
Masai and Lee et al. [52, 57]	AR	Remote Expert	Cartoon	Audio, Visual
Tait et al. [108]	AR	Remote Expert	Cartoon	Audio, Visual
Kurata et al. [49]	AR	Remote Expert	Cartoon	Audio, Visual
Ou et al. [69]	AR	Remote Expert	Cartoon	Audio, Visual
Izadi et al. [45]	AR	Remote Expert	Cartoon	Audio, Visual
Lukosch et al. [55]	AR	Remote Expert	Annotations	Audio, Visual
Gauglitz et al. [25–27]	AR	Remote Expert	Annotations	Audio, Visual
Gupta et al. [30]	AR	Remote Expert	Annotations	Audio, Visual
Zillner et al. [136]	AR	Remote Expert	Annotations	Visual
Utzig et al. [116]	AR	Remote Expert	Annotations	Audio, Visual
Zenati et al. [133–135]	AR	Remote Expert	Annotations	Audio, Visual
Calandra et al. [3]	AR	Remote Expert	Annotations	Audio, Visual
Breakroom [101]	VR	Meeting	Cartoon	Audio, Visual
EngageVR [76]	VR	Meeting	Realistic	Audio, Visual
Glue Collab [39]	VR	Meeting	Realistic	Audio, Visual
MeetInVR [60]	VR	Meeting	Realistic	Audio, Visual
Mozilla Hubs [64]	VR	Meeting	Cartoon	Audio, Visual



Nvidia Holodeck [67]	VR	Meeting	Realistic	Audio, Visual
Stage VR [119]	VR	Meeting	Realistic	Audio, Visual
TechViz VR [110]	VR	Meeting	Realistic	Audio, Visual
TheWild [115]	VR	Meeting	Realistic	Audio, Visual
Vive Sync [106]	VR	Meeting	Cartoon	Audio, Visual
WorldViz [131]	VR	Meeting	Cartoon	Audio, Visual
Regenbrecht et al. [84]	VR	Meeting	Realistic	Audio, Visual
Gu et al. [29]	VR	Meeting	Realistic	Audio, Visual
Schäfer et al. [93]	VR	Meeting	Realistic	Audio, Visual
VRChat [4]	VR	Meeting	Cartoon	Audio, Visual
NeosVR [61]	VR	Meeting	Cartoon	Audio, Visual
Acadicus [118]	VR	Meeting	Cartoon	Audio, Visual
Rumii [100]	VR	Meeting	Realistic	Audio, Visual
VirBELA [117]	VR	Meeting	Realistic	Audio, Visual
Garou [43]	VR	Meeting	Realistic	Audio, Visual
MeetingRoom [59]	VR	Meeting	Cartoon	Audio, Visual
Facebook Horizon [54]	VR	Meeting	Cartoon	Audio, Visual
Second Life [86]	VR	Meeting	Realistic	Audio, Visual
Tan et al. [109]	VR	Meeting	Cartoon	Audio, Visual
Weissker et al. [130]	VR	Meeting	Cartoon	Audio, Visual
Calandra et al. [9]	VR	Meeting	Realistic	Audio, Visual
CollaboVR [34]	VR	Meeting	Realistic	Audio, Visual
IrisVR [44]	VR	Design	Cartoon	Audio, Visual
Hsu et al. [40]	VR	Design	Realistic	Audio, Visual
Lehner et al. [53]	VR	Design	Realistic	Audio, Visual
BigScreen [42]	VR	Event	Realistic	Audio, Visual
Wave [129]	VR	Event	Realistic	Audio, Visual
Sansar [8]	VR	Event	Realistic	Audio, Visual
Higuchi et al. [37]	MR	Meeting	Realistic	Audio, Visual
Speicher et al. [103]	MR	Meeting	Annotations	Audio, Visual
Ryskeldiev et al. [89, 90]	MR	Meeting	Annotations	Visual
Spatial [107]	MR	Meeting	Realistic	Audio, Visual
Regenbrecht et al. [83]	MR	Meeting	Realistic	Audio, Visual, Tactile
Haller et al. [33]	MR	Meeting	Realistic	Audio, Visual, Tactile
Norman et al. [66]	MR	Meeting	Annotations	Audio, Visual
Matthes et al. [58]	MR	Meeting	Realistic	Audio, Visual
Galambos et al. [19, 20]	MR	Meeting	Realistic	Audio, Visual
Bai et al. [2]	MR	Meeting	Cartoon	Audio, Visual
Luxenburger et al. [56]	MR	Meeting	Realistic	Audio, Visual
vTime [120]	MR	Meeting	Realistic	Audio, Visual
PoseMMR [70]	MR	Meeting	Annotations	Audio, Visual
Gamelin et al. [21]	MR	Meeting	Realistic	Audio, Visual
Geollery [14, 15]	MR	Meeting	Cartoon	Audio, Visual
Grønæk et al. [28]	MR	Design	Realistic	Audio, Visual
TeleAR [128]	MR	Design	Cartoon	Audio, Visual, Tactile
Wang et al. [124–127]	MR	Design	Realistic	Audio, Visual, Tactile
Sakong et al. [91]	MR	Design	Realistic	Audio, Visual, Tactile
Sidharta et al. [99]	MR	Design	Realistic	Audio, Visual, Tactile

Ibayashi et al. [41]	MR	Design	Realistic	Audio, Visual, Tactile
Sasikumar et al. [92]	MR	Remote Expert	Cartoon	Audio, Visual
Lee et al. [50, 51]	MR	Remote Expert	Realistic	Audio, Visual
Teo et al. [111–114]	MR	Remote Expert	Realistic	Audio, Visual
Kim et al. [46, 47]	MR	Remote Expert	Annotations	Audio, Visual
Rae et al. [80]	MR	Remote Expert	Realistic	Audio, Visual
Piumsomboon et al. [74, 75]	MR	Remote Expert	Cartoon	Visual
Gao et al. [22, 23]	MR	Remote Expert	Realistic	Audio, Visual
Wang et al. [122]	MR	Remote Expert	Annotations	Audio, Visual
Elvezio et al. [16]	MR	Remote Expert	Realistic	Visual
Pouliquen-Lardy et al. [79]	MR	Remote Expert	Realistic	Audio, Visual
Alem et al. [1]	MR	Remote Expert	Cartoon	Audio, Visual
Higuch et al. [36]	MR	Remote Expert	Realistic	Audio, Visual
Chen et al. [5]	MR	Remote Expert	Annotations	Audio, Visual
Nittala et al. [65]	MR	Remote Expert	Annotations	Audio, Visual, Tactile
Sun et al. [104, 105]	MR	Remote Expert	Cartoon	Audio, Visual
De Pace et al. [12]	MR	Remote Expert	Realistic	Audio, Visual

Table 1. Remote collaboration systems sorted by their respective technology and classified in different categories.

**3.1.1 Category Meeting.** This category is for remote collaboration systems where users share a common workspace or environment for collaboration. These systems usually support media sharing, involve avatars to increase the sense of co-presence, and have interactive elements such as drawing on a whiteboard. In addition, we include use cases with knowledge transfer i.e. educational and learning scenarios in this category. We also include collaborative virtual training in this category.

A collaborative virtual firefighting scenario was implemented by Calandra et al. [9]. The authors implemented a realistic road tunnel simulator for training purposes. Multiple users are able to share a virtual environment in VR and collaborate to extinguish virtual fires and rescue possible victims.

A focus on transferring and obtaining knowledge through augmented and virtual reality remote collaboration systems is shown by Monahan et al. [63], where a web-based VR system for managing and providing educational content online was implemented. The system features an immersive 3D environment, allowing the lecturer to add media and virtual objects. Avatars are able to use gestures e.g. they can raise hands to indicate a question.

Chen et al. [6] created *ImmerTai*, a system which is designed for remote motion training. The participants are able to learn Chinese Taichi in an immersive collaborative environment. Student and a teacher are physically separated and resembled as a full body avatar in the virtual environment. This system includes a motion capturing module utilizing a Microsoft Kinect, transferring the real world motion to their avatars. A motion assessment module is used to rate the movements of the student and give hints for improvement during and after a Taichi session.

Wang et al. [121] use a combination of camera, projector, VR HMD and hand tracker to create a remote collaboration system for knowledge transfer in a manufacturing scenario. A local worker assembles a water pump while a camera is recording and transmitting video footage of the worker's assembling progress to a remote expert. The remote expert views the video material through a VR HMD and transmits visual cues back to the local worker. A projector on the side of the local worker projects the hand movements of the remote expert onto his working surface.



Schäfer et al. [93] used panorama images to create a shared photorealistic virtual environment in a meeting scenario. Users are able to hold virtual presentations with media sharing and hand gestures for interacting with the augmented virtual objects.

Weissker et al. [130] investigated group navigation in virtual immersive environments. The authors implemented a system which allowed users to navigate inside virtual environments together as a group or as individual. In their system, users can attach themselves to others and then organize teleportation movement through the virtual world together. Their results showed advantages in collaborative work when a switch between individual and group navigation is implemented.

**3.1.2 Category Design.** The category design combines remote collaboration systems such as product design [95, 97, 98] and architectural design [7, 38, 40].

One of the earlier works was done by Lehner and DeFanti [53], who used a CAVE system in 1997. CAVE is a 10-foot by 10-foot by 9-foot surround screen which uses projections on the walls and floor. The authors implemented a system which enabled multiple users to share the same environment and discuss vehicle design remotely. The visual representation of other users was achieved by streaming 2D video inside the virtual environment.

Hsu et al. [40] developed an architectural design discussion system with interactive and immersive elements. It features voice communication, object manipulation, mid-air sketching and on-surface sketching. Overall, this tool was implemented to help architects to better understand architecture modeling and to discuss design decisions, even changing models during a remote collaboration session.

The work of Chowdhury et al. [7] implements a collaboration system with an immersive virtual environment specifically created for urban design ideation and generation. The work is intended to be used by non-experts and concludes that even laypeople can take part in the design process of early stage urban design. While one user uses a VR HMD to view, interact and change virtual objects, other participants are able to perceive the changes on a display screen while giving feedback to the VR user.

Hong et al. [38] utilizes the multi-user virtual environment Second Life [86] for creative collaboration with focus on architectural design. They compared the effectiveness of collaborative architectural design between multi user virtual environments and a commercial architectural design software which allows 2D sketching and communication through audio. The authors argue that a remote collaboration system with avatars is more effective than two-dimensional approaches due to shared spatial informations.

Ibayashi et al. [41] created a MR collaboration system which connects users on a tabletop device with a user wearing a VR HMD. The authors use a role based system with designers and occupants. Designers are able to view a 3D environment using a tabletop device. The environment can be changed with a touch interface provided to the designers. The occupant is immersed in this shared 3D environment with a VR HMD and is able to see the changes made by the designers in real-time. A see-through ceiling allows the occupant to see the designers by looking at the ceiling of the 3D environment while the designers are able to see the VR user moving around from the top-view.

A petroleum well planning application was developed by Nittala et al. [65], using hand held devices to augment the surroundings of a remote worker. A local user used a 3D printout, stylus and tablet as an interface to communicate with a remote worker who is on-site coordinating drilling operations. The 3D printout was combined with AR visualisations to provide the local user with an overview of earth's composition near the remote worker. The remote worker is able to see AR annotations made by the local user to plan drilling operations.

**3.1.3 Category Remote Expert.** We identified several systems which include remote collaboration, and use a scenario with a local and a remote user. In such systems, the local user typically executes

a predefined task, being physically present at the target location, while the remote user is generally far away and provides support with instructions or hints. In this type of collaboration scenario, the remote user is often called the remote expert. Many systems are based on a combination of AR and VR [24, 27, 50, 112], while the remote expert typically uses a VR HMD or 2D screen and the local user transmits his surroundings with the help of an AR HMD or a mounted camera.

A mixed reality collaboration system was developed by Piumsomboon et al. [72]. The system enables an AR user to share his local environment with a remote user. It provides collaborative, natural interaction with gaze and hand gesture data transmitted over a network to each user.

Another MR collaboration system was developed by Lee et al. [50]. The authors developed a system in which a host works with an AR HMD mounted with a 360° camera and a guest with a VR HMD. Nonverbal communication cues are transmitted via hand tracking and view awareness. Both users have visual feedback where the counterpart is currently looking at and are able to exchange hand gestures.

Teo et al. [112] introduced a MR remote collaboration system which combines reconstructed scenes obtained through an AR HMD with 360 panorama images. A remote user who receives footage from the AR user can move through the transmitted visual information, without relying on the local user to move. The system has been extended with the functionality that the remote user can trigger a 360 camera with the help of his VR controller to save spherical images that can be accessed independently [114]. The authors propose to add more functionality such as mid-air drawing in 3D to improve the usability of the prototype.

Mixed reality remote collaboration systems supporting local and remote users are especially useful in repairing tasks as Gauglitz et al. [27] suggest. In their work, a remote user is able to see the local user's current view and to annotate the view which is then visible in AR.

An interesting remote collaboration approach is *360Anwywhere*, a framework by Speicher et al [103]. It allows ad-hoc remote collaboration in AR via 360-degree live input. Users are able to add digital annotations by drawing on a 360-degree video stream, either by means of a normal desktop application or mobile devices. The annotations made by remote participants are then visualised at the local physical space through a projector.

Gao et al. [24] implements a mixed reality collaboration system by mounting an RGB and RGB-Depth camera on top of a local user's VR HMD. The VR HMD is used as a Video-See-through device while it captures and transmits its view to the remote counterpart. A RGB-Depth camera is used to obtain a point cloud which is streamed to the remote user. This point cloud is then stitched, enabling an independent view control of the local workspace.

Bai et al. [2] developed a system which supports real-time 3D reconstruction by assembling eight RGB-Depth cameras into one sensor cluster. The remote VR user is able to see the local users surroundings through the transmission of the aligned pointclouds obtained through the RGB-Depth cameras at the local users space. The authors research focus is a shared virtual environment which supports gaze and gesture as visual cues in a remote expert scenario. Although the system supports one-way transmission of natural communication cues only, the results demonstrate advantages by providing natural gaze and gesture cues during collaboration.

Overall, we found that the virtual environment of remote collaboration systems focus on audio-visual stimuli. Some work did not even implement audio, focusing completely on visual feedback [74, 75, 89, 90]. While other work included some tactile feedback, these systems were mostly marker based AR systems and rely markers attached to physical objects [91, 99, 124, 127].

Avatar Type	References
Cartoon	[101] [39] [60] [64] [119] [106] [77, 78] [93] [61] [44] [118] [100] [87] [117] [43] [59] [42] [129] [54] [120] [114] [92] [4] [19, 20] [2] [130] [79] [1] [116]
Realistic	[76] [67] [107] [110] [68] [37] [8] [86] [109] [4] [61] [9] [12] [21] [34]
Full Body	[101] [76] [67] [119] [106] [4] [61] [8] [129] [120] [86] [109] [9] [12] [21]
Head & Hands	[39] [60] [110] [131] [74] [93] [44] [118] [100] [59] [42] [114] [19, 20] [2] [79] [34]
Upper Body	[64] [107] [115] [77, 78] [87] [43] [54] [130]
Reconstructed Model	[76] [107] [110] [68]
Video	[83] [84] [29] [89, 90] [58] [53] [5] [41]
AR annotations	[102] [31, 32] [125] [30] [46] [122] [66] [136] [36] [116] [65] [133, 134]
Hands	[128] [102] [91] [45] [22] [50, 51] [92] [1] [47] [36] [41] [104, 105]
Audio Avatar	[95–98] [99] [25–27] [52, 57] [108] [49] [69] [55] [28] [70] [16] [56] [3]

Table 2. Remote Collaboration Systems classified in avatar categories.

### 3.2 Avatars

Avatars represent entities in virtual environments. We classify the most commonly used avatars in scientific literature as well as in commercial AR/VR/MR software in categories and differentiate them by descriptive terms (see Table 2):

- **Realistic and Cartoon graphics**

We distinguish avatars through visualisation style, i.e. how the 3D model of the avatar is rendered (cartoon or realistic style). One of the reasons we use this as a descriptor is that we want to help researches who are focused on the appearance of avatars. Additionally, there is existing work which addresses certain research questions concerned with avatar visualisation styles. E.g. according to Yoon et al. [132] there is no statistical difference in regards to social presence with different visualisation styles but that user perception differs between cartoon and realistic avatars. A cartoon avatar allows for a more playful atmosphere, whereas realistic avatars tend to represent a professional environment.

- **Avatar Type** We divide avatar types in subcategories: Full Body, Upper Body, Head & Hands and Hands only. Full body avatar refers to humanoid avatars where all limbs are attached to it (e.g. hands, arms, legs etc.). An Upper Body avatar consists of a head, hands and torso but no legs. The Head & Hands type of avatar is composed of a floating head combined with (detached) hands. Hands only means that a user is only represented by virtual hands.
- **Reconstructed Model Avatar** If a system is capable of creating an avatar that resembles the respective user we categorise it as a system that uses *Reconstructed Model* avatars. This category includes avatars that are created from face reconstruction in any form and excludes avatars that do not have a realistic face (e.g. reconstructed/personalized hands only is excluded). Non-reconstructed avatar means in general choosing from existing 3D models, without significant customization options.
- **Video Avatar** Some systems implement avatars as video projections, similar to typical videoconferencing systems. Systems fall under this category if a user is seen as a video feed in an immersive virtual environment. Additionally, we categorise systems which use multiple

cameras to reconstruct 3D video avatars here. As an example, the work of Matthes et al. [58] implements such a system based on multiple depth cameras.

- **Audio Avatar** Although avatars are often represented as a humanoid 3D model, the term avatar is in general used for any kind of user representation in virtual worlds, even including invisible forms. In this work, *Audio Avatar* represents an entity in a system which enables communication with other users, regardless of visual appearance. Systems with this type do not rely on any visual form for users in remote collaboration systems. Users in such systems use audio for communicating with each other.
- **AR Annotations** Systems which use no 3D model for other users but have annotations instead are in this category. This differs from *Audio Avatar* in the sense that *Audio Avatar* uses audio communication only, whereas in *AR Annotations* the remote expert communicates with the local user with annotations, i.e. the other users presence is perceived through visual annotations. Unique avatars are usually not necessary in this scenario, because the roles are clearly separated and the users can distinguish each other by actions. In many systems using *AR Annotation* avatars, audio communication is not implemented. As an example, the work of Zillner et al. [136] uses visual annotations such as text, pictures and freehand drawings to give precise instructions to the local worker without relying on audio communication.

A typical avatar configuration for VR based systems consists of the head (position of VR HMD) and hands (position of controllers) and is usually the most minimalistic avatar in VR scenarios. An avatar consisting solely of virtual hands tends to be used in AR based systems which use hand tracking or gesture techniques [22, 102]. We identified a majority of *Audio Avatars* in AR systems (see Table 2). While VR systems seem to always rely on 3D model representation of other users, AR based approaches often omit visual representation when using remote expert scenarios. In such cases, the local and remote user communicate via audio with each other and share their view [25–27]. Furthermore, some systems rely mainly on visual annotations [31, 31, 102], which we marked as *AR Annotations* in Table 2.

Piumsombon et al. [73] developed a system with an adaptive avatar *Mini-Me* which uses redirected gaze and gestures to enhance remote collaboration with improved social presence. To assess the usefulness of the avatar, a scenario where a remote expert in VR assists a local worker in AR was used. The remote expert was shown to the AR user as a miniature avatar which was able to successfully transmit nonverbal communication cues according to the authors. Although focusing on the novelty of the proposed avatar, the system proved to be useful for overall remote collaboration.

Elvezio et al. [16] developed a system with virtual twins of physical objects. A remote expert uses virtual replicas of physically existing objects to guide a local user performing certain tasks with such objects. In this case, communication with both users only takes place by transmitting the pose of the mentioned objects. In the work of Luxenburger et al. [56] the communication between users takes place through media sharing. A user is filling out a report on a mobile device which is then visible to a remote user by means of a VR HMD.

In the commercial VR remote collaboration system EngageVR [76], users can create their own full body avatar with reconstructed face, by uploading a single picture. Machine learning techniques in the backend of the system reconstruct a fully textured 3D mesh of the head and attaches it automatically to a predefined body models. Some other commercial systems are not as sophisticated and use cartoon like avatars [39, 60, 101]. Other popular systems such as VRChat [4] or NeosVR [61] allow users to create and upload their own avatars. By means of an SDK, they can upload fully animated humanoid avatars regardless of their appearance. The seamless integration of these arbitrary avatars is achieved by applying a specific skeletal structure to the model.

Interactive Feature	References
Shared 3D Object Manipulation	[101] [76] [39] [60] [64] [67] [107] [119] [110] [115] [106] [131] [83] [84] [33] [128] [125] [91] [99] [97] [96] [85] [68] [4] [61] [44] [118] [100] [43] [59] [42] [120] [86] [70] [66] [109] [19, 20] [16] [53] [79] [41] [9] [12] [21] [34] [14, 15]
Media Sharing	[101] [76] [39] [60] [64] [67] [107] [119] [110] [115] [106] [131] [83] [84] [33] [128] [45] [28] [103] [93] [4] [61] [100] [117] [43] [59] [42] [129] [120] [86] [109] [19, 20] [56] [3] [14, 15]
2D Drawing	[76] [39] [60] [64] [107] [119] [131] [83] [33] [29] [128] [69] [45] [28] [103] [37] [61] [100] [59] [19, 20] [3] [34]
Mid-Air Drawing in 3D	[39] [60] [64] [107] [119] [131] [61] [44] [136] [34]
Shared Gaze Awareness	[77, 78] [128] [57] [30] [74, 75] [103] [37] [122] [66] [2] [36] [116]
Convey Facial Expression	[57] [52] [61] [109]
Hand Gestures	[107] [102] [78] [128] [125] [91] [45] [55] [74, 75] [22, 23] [50, 51] [111] [93] [68] [61] [92] [109] [2] [1] [47] [36] [41] [104, 105]
AR annotations	[107] [102] [78] [77] [33] [128] [99] [96–98] [52, 57] [108] [49] [69] [55] [85] [103] [89] [50, 51] [111] [46, 47] [92] [1] [136] [36] [5] [65] [133, 134] [3]
AR viewport sharing	[108] [49] [74, 75] [22, 23] [103] [89, 90] [50, 51] [111, 114] [46] [92] [65] [104, 105] [133, 134] [3]

Table 3. Shared interaction possibilities among remote collaboration systems.

### 3.3 Interaction

We identified common interactive elements in remote collaboration systems that are found in various works and literature. In Table 3 we provide an overview of literature and work which is categorised in multiple different interaction categories. It is to note that a category is not mutually exclusive to another, e.g. a system which uses media sharing might also use hand gestures. This section explains each interaction technique with a few examples. The common features we extracted are as follows:

- (1) Shared 3D Object Manipulation
- (2) Media Sharing
- (3) AR Annotations
- (4) 2D Drawing
- (5) AR Viewport Sharing
- (6) Mid-Air Drawing in 3D
- (7) Hand Gestures
- (8) Shared Gaze Awareness
- (9) Convey Facial Expression

Table 3 guides the reader to interesting and major publications which use the mentioned interaction techniques.

**3.3.1 Shared 3D object Manipulation.** The most commonly shared feature between remote collaboration systems is the possibility to interact and manipulate shared 3D objects in a virtual space.

The type of interaction differs between systems, but the focus is on manipulating one or many 3D objects. AR technology is used by Shen et al. [96–98] where multiple users interact with 3D objects in a collaborative AR environment. A stylus with two markers attached is used as additional interaction tool which enables feature highlighting and 3d object manipulation. More recent work uses hand tracking/gestures to interact with objects [107]. Schlünsen et al. [94] compared free-hand-manipulation with widget-based manipulation techniques. Their study shows that free-hand interaction is preferred over widget-based interaction by users.

**3.3.2 Media Sharing.** Systems which are able to share documents, images, videos and other form of media are categorised here. Haller et al. [33] created a system with a tangible interface, a table with touchscreen for media sharing. It additionally featured sharing media from desktop applications to the tabletop. A web-based VR solution was developed by Monahan et al. [63] which implements media sharing such as videos and images in an educational context.

**3.3.3 AR Annotations.** One of the most commonly used tool for communication and interaction in AR based systems is annotations. Annotation types include 2D drawing, text, or simple pointers. The work of Speicher et al. [103] utilizes a 360 camera to capture the surroundings of one user, while other users are able to draw and annotate on the input stream. The annotations and drawings are then visualised by a projector to the local user’s physical space. Kurata et al. [49] present a wearable HMD which receives remotely annotated input in form of drawings. A special feature of this system is a laser pointer that enables the wearer of an HMD to draw the attention of remote users towards a certain object by pointing on it.

**3.3.4 2D Drawing.** Especially in remote collaboration systems with focus on replicating a virtual meeting scenario, drawing on surfaces is a widely used feature. In more sophisticated systems such as Glue Collaboration Platform [39], users are able to place a virtual whiteboard. This whiteboard can be re-positioned and resized allowing multiple users to draw with virtual pens in many sizes and colors. Mimicking real world objects, it is also possible to use an eraser.

**3.3.5 AR viewport sharing.** This category includes systems which implemented sharing a user’s view perspective. Tait et al. [108] created a system which reconstructs a local user’s environment by using depth sensors attached to an HMD. Reconstructing the environment from a local scene, remote users can move independently through the virtual environment. The local user is then represented as a frustum in the reconstructed scene, allowing the remote user to see where the local counterpart is looking at. The study of Tait et al. [108] suggests that implementing view independence between local and remote user improves task completion time. Sasikumar et al. [92] combined AR and VR users together and enabled view frustum sharing which is visible to the local user as a grey cuboid. The goal of their work was to convey nonverbal communication cues such as eye gaze and hand gestures.

**3.3.6 Mid-Air Drawing in 3D.** Mid-Air drawing allows users to create 3D paintings, which can then be observed from multiple users in different angles. This interaction method seems to be mostly available in commercial systems such as Glue Collab [39] or [64]. A user is able to draw in the air of the virtual world by utilizing a VR controller as virtual pen. Other systems such as the one proposed by Zillner et al. [136] implement a remote expert scenario, where one user is streaming his surroundings with an RGB-D camera and a remote expert is observing and annotating for assistance. The remote expert is able to segment objects, to create animations, to draw on geometry and to place annotations which can be viewed by the AR user.

**3.3.7 Hand Gestures.** Systems which utilize hand gestures and convey hand movements through the remote collaborative space are included in this category. Sophistated systems such as *Spatial*



[107] use AR HMD's to enable a full interaction with the 3D environment via a hand tracker. Tan et al. [109] implemented a VR telepresence system which allowed multiple users to interact with objects and watch videos together. The authors used motion capture gloves to animate arms, hands and fingers of a VR avatar. Kim et al. [47] implemented a MR collaboration system to evaluate combinations of visual communication cues using gestures. The authors found that certain combinations of communication cues such as *hand visualisation* together with *finger pointing direction* does not provide any significant benefit for remote collaboration.

**3.3.8 Shared Gaze Awareness.** Systems which allow users to share gaze awareness belong into this category. We include systems that allow precise tracking and transmission of gaze awareness and exclude systems that indicate gaze perception by head rotation only. E.g. Galombos et al. [20] is not included since the gaze direction of a user is only indicated by the direction the avatar is facing. As an example, Speicher et al. [103] created a system that allows the participants to show exactly which position they are looking at in a 360° video feed. The work of Poppe et al. [77, 78] uses avatars around a virtually augmented table and positions them according to the gaze information of the corresponding user. Billinghurst et al. developed *Empathy Glasses* [52, 57] which is a HMD that enables streaming a live video feed with accurate gaze information. Norman et al. [66] implemented a system which gives direct visual feedback of other user's gaze behavior. Using a system that combines multiple AR HMD's with a desktop PC, participants are asked to place virtual furniture on a regular table in a collaborative manner.

**3.3.9 Convey Facial Expression.** This category addresses work which is able to transmit facial expressions to other users. Systems which use video transmission of other user's faces are not included. Lee [52] and Masai[57] et al. use *Empathy Glasses* to transmit facial expressions bidirectionally. A local user's face is analyzed by the built in modules of the glasses, while a remote user's expression is tracked via webcam. The authors of [109] integrated lip syncing into their VR remote telepresence system in order to enable a more immersive communication experience.

## 4 PROFESSIONAL MEETINGS THROUGH VIRTUAL AND AUGMENTED REMOTE COLLABORATION

Currently, there are already many tools and collaboration systems available which utilize virtual or augmented reality. In this section we consider commercial and professional systems which support more than 10 users simultaneously [39, 60, 64, 67, 76, 101, 107, 110, 115, 119, 131]. During the COVID-19 outbreak in early 2020, one of the biggest scientific conferences for virtual reality research (IEEE VR 2020) was held completely online and virtual. To this date, it was the first major conference held completely virtually. During this time, the organizing committee faced a very challenging task to provide a pleasant conference experience for all participants without the need for a physical presence. During this precedent case, the whole conference featured livestreams for each track, where each presenter had the opportunity to either give a presentation with pre-recorded video, live video transmission or presenting with a VR HMD. Additionally, there were multiple virtual meeting rooms in which participants could join and then interact and network with other users using VR HMD's. The virtual meeting rooms utilized during this conference were built on Mozilla Hubs [64]. It features a web-based meeting room creation software, which enables users to develop and maintain their own meeting experience. Utilizing a cartoon graphics style, this system works with a desktop application, internet browser and even mobile devices, covering a broad possible user audience. Avatars are chosen from existing, pre-defined models. They use a mix between an abstract representation and Upper Body avatars by using a robot-like representation for users.

While this solution uses an open source approach, there are several commercial products available [39, 60, 101, 115, 119] which are similar in terms of remote collaboration utilizing VR technology. *Glue Platform* [39] is a platform for business professionals offering immersive 3D graphics in cartoon style. It is built to be used with virtual reality HMD's and claims to be an extension to the everyday working life. Main features include spatial audio, 3D avatars, interactive and persistent objects. The avatars use a simple Head & Hands approach.

Another commercial software available is called Breakroom [101]. It supports VR HMD's, is available for multiple platforms, and features full body avatars.

#### 4.1 Comparison of commercial systems

In this section, we compare existing commercial/professional remote collaboration systems which are already in use by industry and individuals. We found several VR remote collaboration systems which are used in a professional context [39, 67, 76, 107]. These systems have many common aspects, such as allowing many users to participate, or the collaboration tools available in the virtual environment. Commercial systems seem to differ mostly in terms of avatars, environments and visualisation styles. Especially in interaction, the systems share many common collaboration tools inside the virtual world. A typical way to collaborate in these systems is using some drawing mechanism or mid-air drawing in 3D space. Copying real world interaction methods, users can sketch on virtual black- or whiteboards and share their results in real time with each other. Some of the mentioned systems also implement mid-air drawing, and allow sketches to be observed from multiple angles. Usually there is also a name indicator, displaying the name of participants. This seems to be necessary even in systems with completely reconstructed or personalised avatars such as EngageVR [76] and Spatial [107]. A common use case in the aforementioned systems is sharing and observing virtual objects together. Systems such as EngageVR [76] or Neos VR [61] allow users to place any 3D object previously added to a catalogue. In some cases it is also possible to show other users information by floating markers i.e. annotations. The virtual environment in these systems usually have a table and multiple seats to copy the physical space of real world meetings. A common scenario involves users to sit on virtual chairs and to present on a virtual tv or projector. Some more advanced systems (e.g. Mozilla Hubs [61, 64]) allow screen sharing to the virtual world. Even more sophisticated systems (e.g. EngageVR [76]) extends screen sharing functionality with full control over normal desktop applications in the virtual world. Additionally, most developers are implementing platform independence, supporting devices such as desktop, VR HMD's, tablets and smartphones. Mozilla Hubs [64] is fully available in a web browser.

#### 4.2 Exploring the strengths and weaknesses of commercial remote collaboration systems

As part of this survey, we identified strengths and weaknesses of popular commercial virtual meeting systems. This survey includes a total of 25 commercial and professional systems (24 VR and one MR based system). The combined strength and weaknesses of these systems include:

- + Many users are able to join and participate in virtual meetings simultaneously [60, 64, 101], usually allowing about 20-50 people to share a virtual space simultaneously.
- + Intuitive interaction possibilities such as mid-air drawing [64], drawing on white- or blackboards [60], media sharing and screen sharing [76].
- + Spatial audio which enables localisation of an audio source during meetings more naturally [39].
- + Persistent virtual objects which exist through multiple sessions (e.g. a drawing from a session before is still present in the next session) [39].

- + Placing arbitrary 3D objects in a shared virtual environment [61, 76].
- + Reconstructed, personalised avatars [76] and user created avatars through an API which is provided by the developers [4, 61].
- + Availability on multiple platforms: desktop, mobile devices and web browser [39, 64].
- Some systems are enterprise solutions [39, 67], tailored to the specific needs of companies, what renders them unattractive or even inaccessible to the general public.
- One major issue with these systems is the missing transmission of nonverbal communication cues to the virtual environment by means of an avatar, which is an important feature of traditional face-to-face collaboration.
- Another weak point of the current systems is dynamic content creation for the virtual environment. These systems are limited to the choice of virtual environments provided by the developer or need expert knowledge to create them [39, 60, 64, 76].

Some systems implement a whole Metaverse that acts as a virtual space for many scenarios such as training, collaboration and other social activities inside a virtual world. NeosVR [61] and VRChat [4] allow experienced users to create custom environments, avatars and interactive objects by providing an API within the game engine in which the platform is implemented.

Overall, the landscape of commercial and professional systems is vast but many systems excel in certain aspects but lack novelty in others. As an example, the immersive remote collaboration software VirBELA [117] allows hundreds of users to participate in a virtual world simultaneously but the avatars lack personalisation. This means that users should think carefully about the requirements a system has to meet and then decide which system to use.

## 5 DISCUSSION AND SURVEY RESULT

In this section we provide an overview of the insights and statistical data we gathered throughout the survey. Especially, we want to emphasize on our taxonomy consisting of *Environment*, *Avatars* and *Interaction* and important findings in each category. It is to note that we included professional and commercial systems in our survey (24 VR based systems and one MR based system). This implies that some of the discussed applications are not published in scientific articles and the implementation of interactive features, virtual environment and the audiovisual representation of users is subject to change in the future. For example, a discussed system does not implement full body avatars at the time of writing, but could implement it later on.

### 5.1 Environment

We categorised remote collaboration systems according to their technology: AR/VR/MR, use case, visualisation styles and sensory inputs. The technology distribution of included systems is shown in Figure 4.

**5.1.1 Technology.** In the category of VR based synchronous remote collaboration systems we included 24 commercial and 9 research oriented systems. The majority of systems with VR technology are commercial systems, which could be an indicator that VR based systems are currently more under development in the industrial sector rather than the research community and therefore could be placed on the plateau of productivity. 18 purely AR based systems were found in which users are able to communicate and collaborate in real-time by means of Video-See-through, Optical-See-through AR or a combination of both. To the best of our knowledge there is no commercial system which is solely based on AR allowing real-time synchronous remote collaboration. MR based systems form the majority of the discussed systems. 35 research oriented and one commercial system are included in this category.

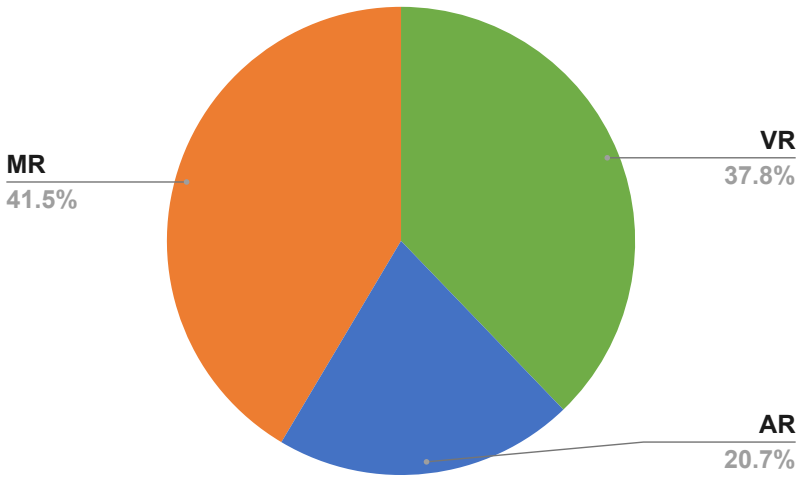


Fig. 4. Distribution of used technology in the discussed systems.

**5.1.2 Use cases.** We divided systems into three different main use cases: *Meeting*, *Design* and *Remote Expert* as they are most popular in the literature. The distribution of mentioned use cases with respect to their technology is shown as a graph in Figure 5.

Systems based on VR technology often involve many users (more than two), with a focus on Meeting and Design use cases. The virtual environments used in VR technology tend to involve all participants equally, i.e. they can see the same things and have the same input modalities, whereas MR based systems often have asymmetric inputs. For example, an asymmetric input method would be when two users share a virtual environment and one has a keyboard and the other has a VR headset with controllers as input device. There was no purely VR based system which implemented a *Remote Expert* scenario although it is by far the most popular use case in AR based systems. While VR systems have a focus on *Meeting* scenarios and AR systems a focus on *Remote Expert* scenarios, MR systems are more distributed throughout use cases. This indicates a strong correlation between the hardware and its respective benefits in certain use cases. For example, VR HMD's are more beneficial in meeting scenarios with an immersive, shared virtual environment, while AR HMDs have more advantages in supporting a user with a remote expert. Since MR is a mix between AR and VR, the distribution of cases is more equally.

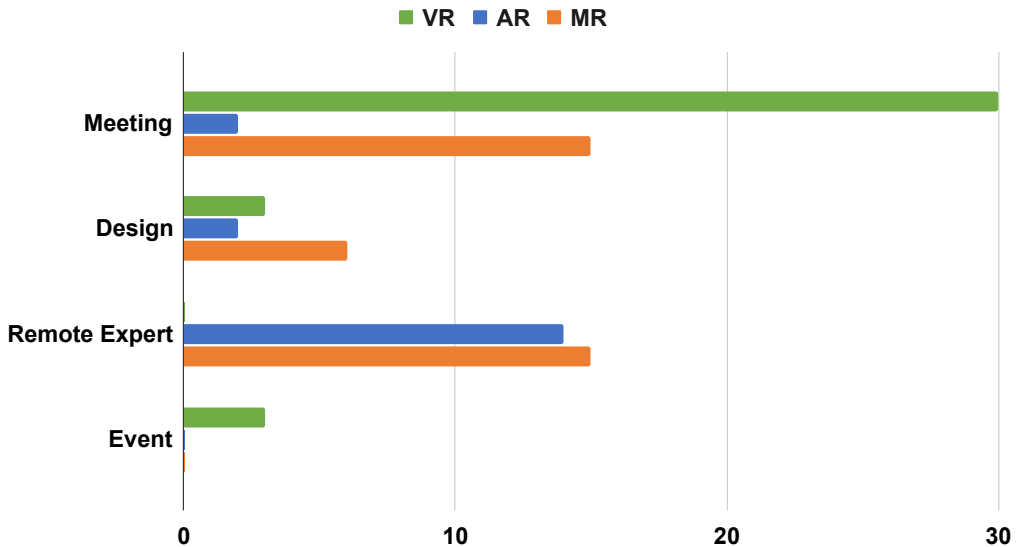


Fig. 5. Distribution of use cases in the discussed systems.

**5.1.3 Visualisation Styles.** We found no evidence that there is a clear divergence in the visualisation styles of synchronous remote collaboration systems. We differentiated between cartoon, realistic and annotation style. A system was labeled cartoon style when there was an obvious simplification of 3D objects with stylized visualisation. Older systems with realistic visualisation are also counted here, as long as no stylized visualisation technique was used. Many AR based systems use annotated video only while some use rendered 3D objects with a cartoon style.

**5.1.4 Sensory Impressions.** The recent literature seems to focus more on audiovisual systems and does not support other sensory impressions such as olfactory and tactile cues. Although tangible interfaces were popular especially in AR technology, the focus is drifting towards audiovisual systems. One of the reasons might be that markers are no longer required to be placed on physical objects which often already implied a tangible system, if markers are placed on non-stationary objects. Additionally, the tracking accuracy of AR systems is constantly being improved. An example is Rambach et al. [81, 82], where the authors use SLAM technology for accurate object tracking without markers.

## 5.2 Avatars

The systems were analysed with respect to their specific avatar implementation. An explanation of our avatar categorisation is found in section 3.2. We found that there is not a single most used avatar type for AR, VR or MR systems. However, we have seen that certain types of avatars are not used in combination with particular technologies. An overall distribution of avatars in the discussed systems is shown in Figure 6.

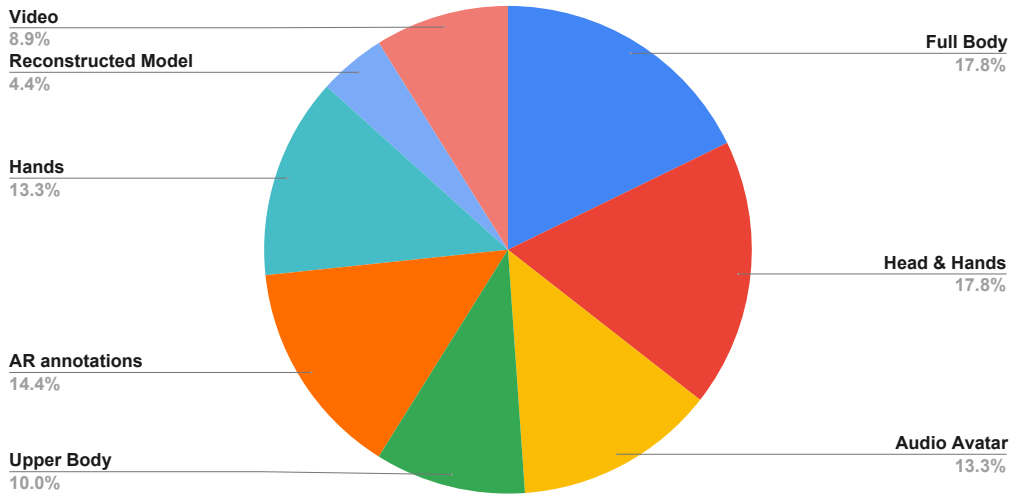


Fig. 6. Different avatar types in presented systems.

Filtering the avatar types by technology gives interesting insights into the spectrum of personal embodiment in synchronous remote collaboration systems. Looking at Figure 7 we found that *Full Body*, *Upper Body* and *UHead & Hands* are most prominent in VR based systems. To the best of our knowledge, there is no AR system that uses a full body 3D model for representation of other users. The closest approach to a full body 3D model representation in an AR system is Holoportation [68], but since several RGB-Depth cameras are involved that transmit video in real-time, we classify this approach as a video-based avatar system. The *Hands* approach for avatars is mostly used in AR and MR systems. This is due to the fact that hand gestures provide a natural and easily understood way to convey visual communication cues in a *Remote Expert* scenario, which is used in AR and MR based systems. Overall we found a lack of systems which utilizes *Reconstructed Models*, i.e. an avatar that is created through face reconstruction in the virtual environment. A reason for this is that the technology for easy reconstruction of humans is not yet widely available for researchers and industry implementing such systems. Not surprisingly, *Audio Avatars* are mostly used in AR and MR based systems. One of the reasons is that the shared virtual environment is formed through video transmission and the communication with other users happens through audio. VR systems use a rendered virtual 3D scene that simplifies finding other users in a virtual environment, allowing for a more sophisticated visualisation approach such as *Full Body* or other types of visual avatars.

### 5.3 Interaction

Examining the different types of interaction in remote collaboration systems, a general majority of the interaction type *Shared Object Manipulation* is found. A possible explanation for this is the ease of implementation and general versatility of a task involving the manipulation of 3D objects. Another popular feature is *Media Sharing*, i.e. possibility to share images, videos and other form of media. Some interaction types are more popular in VR technology, such as *Shared Object Manipulation*, *Media Sharing*, *2D Drawing* and *Mid-Air Drawing in 3D*. The interaction types *Viewport Sharing*, *AR annotations* and *Shared Gaze Awareness* had no implementations at all in a pure VR scenario. Transmitting facial expressions by using avatars is by far the least prominent feature in remote collaboration systems, even though it is an important step towards more natural



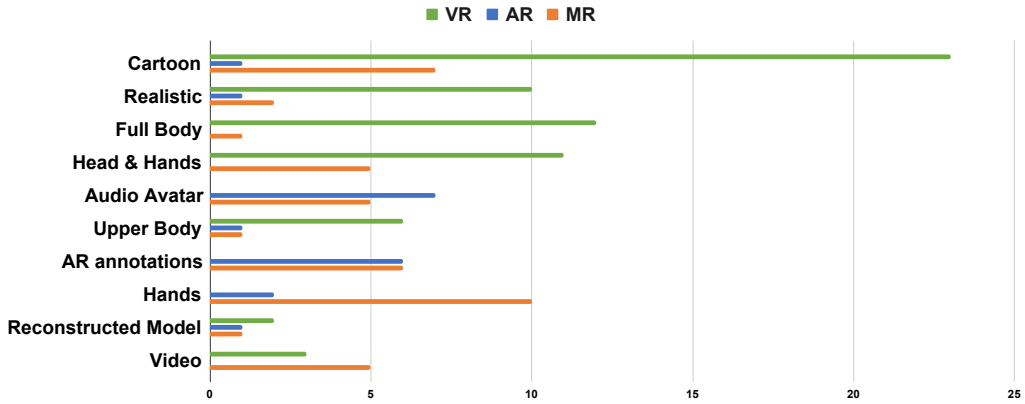


Fig. 7. Different avatar types in respect to their used technology.

conversations over distance. A comparative graph about the interaction types and their presence in the discussed work is shown in Figure 8.

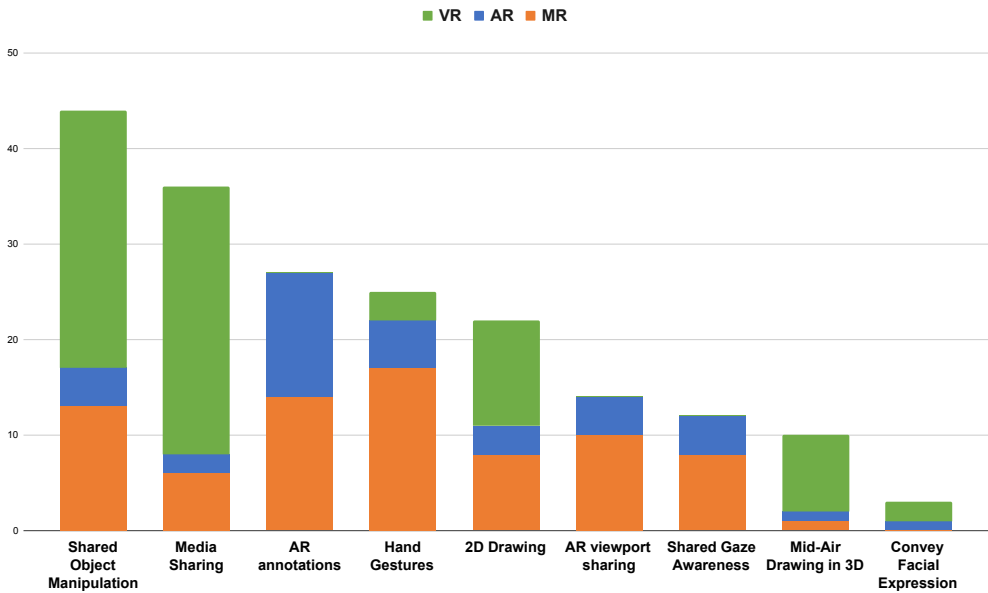


Fig. 8. Interaction Types in respect to the used technology.

Additionally, we found that most of the commercial remote collaboration systems rely solely on VR technology (24 VR systems and one MR system). Since the purchase of commercially available VR HMD’s and a reasonable PC is required to operate these systems, their actual use is still limited. Therefore most professional and commercial systems tend to implement a desktop and mobile

version of the VR application. Only a few companies, such as Spatial [107] focus on integrating AR, VR, desktop and mobile together. Deploying the same application on multiple hardware devices, each with different input modalities, raises new issues such as asymmetric input. For example, a mobile device user can most likely collaborate only by voice or minimal interaction within a virtual environment, compared to a user with a VR HMD and full body tracking. In an attempt to solve the asymmetric input problem, Fleury et al. [18] investigated how the same interaction possibilities can be realized in virtual space, in a CAVE environment, and a high-resolution 2D display in wall format. Another approach is taken by Pouliquen-Lardy et al. [79], which implemented a multi-user remote collaboration MR system to study the asymmetric effects of different input modalities. The authors used an approach with two different roles, a guide who could observe and communicate via audio, and a manipulator who was able to manipulate a 3D object. The overall result of their study suggests that it is not necessary to develop symmetric interaction for all users during remote collaboration, but rather the same interaction possibilities for roles. For example, a user in the role of a guide should be able to observe and communicate via audio regardless of the hardware they are using.

## 6 CONCLUSION

During this survey, we identified different kinds of virtual environments, various types of avatars and many shared interactive elements for synchronous remote collaboration systems. To help researchers in various research fields we created a taxonomy consisting of *Environment*, *Interaction* and *Avatars*. This approach aims to provide condensed information to specific topics which needs to be addressed while designing and implementing a synchronous remote collaboration system.

In the category *Environment* it is found that the most prominent scenario for VR based systems is *Meeting*. While the *Design* scenario is rather equally distributed among the different technologies, *Remote Expert* is dominated by the use of AR and MR technology. Regarding *Avatars*, most VR systems use a rather cartoonish visualisation style with representations ranging from full body to head & hands equally distributed. Contrary, AR and MR systems tend to have user representations with hands only, use annotations, or rely on audio communication and omit visual avatars. Regarding the different *Interaction* types, the most prominent features are shared object manipulation followed by media sharing.

Overall it can be said that congresses, conferences, conventions, and other events get reasonable alternatives to in person meetings with synchronous AR/VR/MR remote collaboration systems. However, systems based on AR, VR, or MR technology rely on different strengths which are reflected in this overview. For example, it appears that AR based systems focus on sharing a person's surroundings with emphasis on stimulating the audiovisual senses. This has recently led to a focus on remote and local user scenarios in AR/MR systems. In such scenarios, typically a user is physically present at a certain location and shares his/her environment, which is then perceived by other users as a virtual world that can be augmented with virtual objects. This means however, that each type of user (local user or remote expert) has different input and output options in their respective virtual environment. These asynchronous input modalities often pose problems and therefore this subject is being actively researched by the scientific community. As an example, a local user streams the surroundings with an AR HMD, while a remote user can add annotations by using e.g. a VR HMD or tablet device, which defines separate roles during remote collaboration.

In contrast, VR based systems tend to involve all participants equally, allowing each user to participate in the same virtual environment and the same communication and collaboration possibilities. VR systems are also commonly used to bring large events online. We also noticed a trend in VR technology to focus more on design and meetings rather than remote expert scenarios, as is mainly the case with AR and MR systems. Independently between AR, VR and MR the research

focus in remote collaboration software drifts towards integrating non-verbal communication cues. Some researchers focus on developing solutions for nonverbal communication transmission, but this has not yet been integrated in professional and commercial remote collaboration systems. This research includes Wang et al. [128] focusing on eye-contact, Masai et al. [57] conveying facial expressions and Lei et al. [24] implementing mutual hand gestures to name a few.

Since novel remote collaboration systems are being developed on a continuous basis, it can be expected that systems based on AR/VR/MR technology will be integrated into our daily life. However, until this is the case, more research is required. VR-based systems require more comfortable hardware to allow a daily usage. Avatars within such systems must transmit the non-verbal communication signals of the participants in order to get closer to a face to face meeting. AR-based systems are still in need of better tracking algorithms and cheaper hardware so that these systems can be used by as many people as possible in everyday life.

## ACKNOWLEDGMENTS

This work was partially supported by Offene Digitalisierungsallianz Pfalz, BMBF, grant number 03IHS075B and the EU Research and Innovation programme Horizon 2020 under the grant agreement ID: 883293.

## REFERENCES

- [1] Leila Alem, Franco Tecchia, and Weidong Huang. 2011. HandsOnVideo: Towards a gesture based mobile AR system for remote collaboration. In *Recent trends of mobile collaborative augmented reality systems*. Springer, 135–148.
- [2] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [3] Davide Calandra, Alberto Cannavò, and Fabrizio Lamberti. 2021. Improving AR-powered remote assistance: a new approach aimed to foster operator’s autonomy and optimize the use of skilled resources. *The International Journal of Advanced Manufacturing Technology* (2021), 1–18.
- [4] VR Chat. 2018. Create and Play in Virtual Worlds. *VRChat Inc* (2018).
- [5] Henry Chen, Austin S Lee, Mark Swift, and John C Tang. 2015. 3D collaboration method over HoloLens™ and Skype™ end points. In *Proceedings of the 3rd International Workshop on Immersive Media Experiences*. 27–30.
- [6] Xiaoming Chen, Zhibo Chen, Ye Li, Tianyu He, Junhui Hou, Sen Liu, and Ying He. 2019. ImmerTai: Immersive Motion Learning in VR Environments. *Journal of Visual Communication and Image Representation* 58 (2019), 416–427.
- [7] Shuva Chowdhury and Marc Aurel Schnabel. 2019. Laypeople’s Collaborative Immersive Virtual Reality Design Discourse in Neighborhood Design. *Frontiers in Robotics and AI* 6 (2019), 97.
- [8] Wookey Project Corp. 2018. Sansar. <https://www.sansar.com/>
- [9] Calandra Davide, Gabriele Praticco Filippo, Migliorini Massimo, Verda Vittorio, and Fabrizio Lamberti. 2021. A Multi-role, Multi-user, Multi-technology Virtual Reality-based Road Tunnel Fire Simulator for Training Purposes. In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2021)*. ScitePress, 96–105.
- [10] Ryan Anthony J de Belen, Huyen Nguyen, Daniel Filonik, Dennis Del Favero, and Tomasz Bednarz. 2019. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering* 3, 2 (2019), 181.
- [11] Francesco De Pace, Federico Manuri, Andrea Sanna, and Claudio Fornaro. 2020. A systematic review of Augmented Reality interfaces for collaborative industrial robots. *Computers & Industrial Engineering* 149 (2020), 106806.
- [12] Francesco De Pace, Federico Manuri, Andrea Sanna, and Davide Zappia. 2019. A comparison between two different approaches for a collaborative mixed-virtual environment in industrial maintenance. *Frontiers in Robotics and AI* 6 (2019), 18.
- [13] Huong Q Dinh, Neff Walker, Larry F Hodges, Chang Song, and Akira Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, 222–228.
- [14] Ruofei Du, David Li, and Amitabh Varshney. 2019. Geollery: A Mixed Reality Social Media Platform. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI)*. ACM, 13. <https://doi.org/10.1145/3290605.3300915>

- [15] Ruofei Du, David Li, and Amitabh Varshney. 2019. Project Geollery.com: Reconstructing a Live Mirrored World With Geotagged Social Media. In *Proceedings of the 24th International Conference on Web3D Technology (Web3D)*. ACM, 1–9. <https://doi.org/10.1145/3329714.3338126>
- [16] Carmine Elvezio, Mengu Sukan, Ohan Oda, Steven Feiner, and Barbara Tversky. 2017. Remote collaboration in AR and VR using virtual replicas. In *ACM SIGGRAPH 2017 VR Village*. 1–2.
- [17] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131 (2019), 81–98.
- [18] Cédric Fleury, Nicolas Ferey, Jean-Marc Vézien, and Patrick Bourdot. 2015. Remote collaboration across heterogeneous large interactive spaces. In *2015 IEEE Second VR international workshop on collaborative virtual environments (3DCVE)*. IEEE, 9–10.
- [19] Péter Galambos, Péter Baranyi, and Imre J Rudas. 2014. Merged physical and virtual reality in collaborative virtual workspaces: The VirCA approach. In *IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 2585–2590.
- [20] Péter Galambos, Ádám Csapó, Péter Zentay, István Marcell Fülöp, Tamás Haidegger, Péter Baranyi, and Imre J Rudas. 2015. Design, programming and orchestration of heterogeneous manufacturing systems through VR-powered remote collaboration. *Robotics and Computer-Integrated Manufacturing* 33 (2015), 68–77.
- [21] Guillaume Gamelin, Amine Chellali, Samia Cheikh, Aylene Ricca, Cedric Dumas, and Samir Otmane. 2020. Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. *Personal and Ubiquitous Computing* (2020), 1–18.
- [22] Lei Gao, Huidong Bai, Gun Lee, and Mark Billinghurst. 2016. An oriented point-cloud view for MR remote collaboration. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*. 1–4.
- [23] Lei Gao, Huidong Bai, Rob Lindeman, and Mark Billinghurst. 2017. Static local environment capturing and sharing for MR remote collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*. 1–6.
- [24] Lei Gao, Huidong Bai, Thammathip Piumsomboon, G Lee, Robert W Lindeman, and Mark Billinghurst. 2017. Real-time visual representations for mixed reality remote collaboration. (2017).
- [25] Steffen Gauglitz, Cha Lee, Matthew Turk, and Tobias Höllerer. 2012. Integrating the physical environment into mobile remote collaboration. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*. 241–250.
- [26] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*. 197–205.
- [27] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. World-stabilized annotations and virtual scene navigation for remote collaboration. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 449–459.
- [28] Kaj Grønbaek, Kristian Kroyer Gundersen, Preben Mogensen, and Peter Ørbæk. 2001. Interactive Room Support for Complex and Distributed Design Projects.. In *INTERACT*. 407–414.
- [29] Ning Gu, Mi Jeong Kim, and Mary Lou Maher. 2011. Technological advancements in synchronous collaboration: The effect of 3D virtual worlds and tangible user interfaces on architectural design. *Automation in Construction* 20, 3 (2011), 270–278.
- [30] Kunal Gupta, Gun A Lee, and Mark Billinghurst. 2016. Do you see what I see? The effect of gaze tracking on task space remote collaboration. *IEEE transactions on visualization and computer graphics* 22, 11 (2016), 2413–2422.
- [31] Pavel Gurevich, Joel Lanir, and Benjamin Cohen. 2015. Design and implementation of teleadvisor: a projection-based augmented reality system for remote collaboration. *Computer Supported Cooperative Work (CSCW)* 24, 6 (2015), 527–562.
- [32] Pavel Gurevich, Joel Lanir, Benjamin Cohen, and Ran Stone. 2012. TeleAdvisor: a versatile augmented reality tool for remote assistance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 619–622.
- [33] Michael Haller, Mark Billinghurst, J Leithinger, D Leitner, and Thomas Seifried. 2005. Coeno: enhancing face-to-face collaboration. In *Proceedings of the 2005 international conference on Augmented tele-existence*. 40–47.
- [34] Zhenyi He, Ruofei Du, and Ken Perlin. 2020. CollaboVR: A Reconfigurable Framework for Multi-User to Communicate in Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 12. <https://doi.org/10.1109/ISMAR50242.2020.00082>
- [35] Morton L Heilig. 1962. Sensorama simulator. US Patent 3,050,870.
- [36] Keita Higuch, Ryo Yonetani, and Yoichi Sato. 2016. Can eye help you? Effects of visualizing eye fixations on remote collaboration scenarios for physical tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 5180–5190.

- [37] Keita Higuchi, Yinpeng Chen, Philip A Chou, Zhengyou Zhang, and Zicheng Liu. 2015. Immerseboard: Immersive telepresence experience using a digital whiteboard. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2383–2392.
- [38] Seung Wan Hong, Ahmed El Antably, and Yehuda E Kalay. 2019. Architectural design creativity in Multi-User Virtual Environment: A comparative analysis between remote collaboration media. *Environment and Planning B: Urban Analytics and City Science* 46, 5 (2019), 826–844.
- [39] Glue HQ. 2020. Glue Universal Collaboration Platform. <https://glue.work/>
- [40] Ting-Wei Hsu, Ming-Han Tsai, Sabarish V Babu, Pei-Hsien Hsu, Hsuan-Ming Chang, Wen-Chieh Lin, and Jung-Hong Chuang. [n.d.]. Design and Initial Evaluation of a VR based Immersive and Interactive Architectural Design Discussion System. ([n. d.]).
- [41] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. 2015. Dollhouse vr: a multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Emerging Technologies*. 1–2.
- [42] Bigscreen inc. 2020. Big Screen. <https://www.bigscreenvr.com/>
- [43] Garou Inc. 2020. Garou. <https://www.garou.io/>
- [44] IrisVR Inc. 2019. IrisVR. <https://irisvr.com/>
- [45] Shahram Izadi, Ankur Agarwal, Antonio Criminisi, John Winn, Andrew Blake, and Andrew Fitzgibbon. 2007. C-Slate: a multi-touch and object recognition system for remote collaboration using horizontal surfaces. In *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*. IEEE, 3–10.
- [46] Seungwon Kim, Mark Billinghurst, and Gun Lee. 2018. The effect of collaboration styles and view independence on video-mediated remote collaboration. *Computer Supported Cooperative Work (CSCW)* 27, 3-6 (2018), 569–607.
- [47] Seungwon Kim, Gun Lee, Weidong Huang, Hayun Kim, Woontack Woo, and Mark Billinghurst. 2019. Evaluating the combination of visual communication cues for HMD-based mixed reality remote collaboration. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [48] Sanjeev Kumar, Jatin Chhugani, Changkyu Kim, Daehyun Kim, Anthony Nguyen, Pradeep Dubey, Christian Bienia, and Youngmin Kim. 2008. Second life and the new generation of virtual worlds. *Computer* 41, 9 (2008), 46–53.
- [49] Takeshi Kurata, Nobuchika Sakata, Masakatsu Kouroggi, Hideaki Kuzuoka, and Mark Billinghurst. 2004. Remote collaboration using a shoulder-worn active camera/laser. In *Eighth international symposium on wearable computers*, Vol. 1. IEEE, 62–69.
- [50] Gun A Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. 2017. Sharedsphere: MR collaboration through shared live panorama. In *SIGGRAPH Asia 2017 Emerging Technologies*. 1–2.
- [51] Gun A Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. 2018. A user study on mr remote collaboration using live 360 video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 153–164.
- [52] Youngho Lee, Katsutoshi Masai, Kai Kunze, Maki Sugimoto, and Mark Billinghurst. 2016. A remote collaboration system with empathy glasses. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, 342–343.
- [53] Valerie D Lehner and Thomas A DeFanti. 1997. Distributed virtual reality: Supporting remote collaboration in vehicle design. *IEEE Computer Graphics and Applications* 17, 2 (1997), 13–17.
- [54] Facebook Technologies LLC. 2020. Facebook Horizon. <https://www.oculus.com/facebookhorizon/>
- [55] SG Lukosch, Ronald Poelman, Oytun Akman, and PP Jonker. 2012. A novel gesture-based interface for crime scene investigation in mediated reality. In *Workshop at CSCW 2012: Exploring collaboration in challenges environments: From the car to the factory and beyond, 11-15 februari 2012, Seattle, Washington, USA*. ACM.
- [56] Andreas Luxenburger, Alexander Prange, Mohammad Mehdi Moniri, and Daniel Sonntag. 2016. MedicaLVR: towards medical remote collaboration using virtual reality. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 321–324.
- [57] Katsutoshi Masai, Kai Kunze, Maki Sugimoto, and Mark Billinghurst. 2016. Empathy glasses. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 1257–1263.
- [58] Carl Matthes, Tim Weissker, Emmanouil Angelidis, Alexander Kulik, Stephan Beck, Andre Kunert, Anton Frolov, Sandro Weber, Adrian Kreskowski, and Bernd Froehlich. 2019. The Collaborative Virtual Reality Neurorobotics Lab. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1671–1674.
- [59] MeetingRoom.io. 2020. MeetingRoom. <https://meetingroom.io/>
- [60] MeetInVR. 2020. MeetInVR. <https://meetinvr.net/>
- [61] Neos VR Metaverse. 2020. NeosVR. <https://neosvr.com/>
- [62] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, Vol. 2351. International Society for Optics and Photonics, 282–292.

- [63] Teresa Monahan, Gavin McArdle, and Michela Bertolotto. 2008. Virtual reality for collaborative e-learning. *Computers & Education* 50, 4 (2008), 1339–1353.
- [64] Mozilla. 2020. Hubs by Mozilla. <https://hubs.mozilla.com/>
- [65] Aditya Shekhar Nittala, Nico Li, Stephen Cartwright, Kazuki Takashima, Ehud Sharlin, and Mario Costa Sousa. 2015. PLANWELL: spatial user interface for collaborative petroleum well-planning. In *SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications*. 1–8.
- [66] Mitchell Norman, Gun A Lee, Ross T Smith, and Mark Billinghurst. 2019. The impact of remote user’s role in a mixed reality mixed presence system. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*. 1–9.
- [67] NVIDIA. 2020. NVIDIA Holodeck. <https://nvidia.com/holodeck/>
- [68] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L Davidson, Sameh Khamis, Mingsong Dou, et al. 2016. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. 741–754.
- [69] Jiazhi Ou, Xilin Chen, Susan R Fussell, and Jie Yang. 2003. DOVE: Drawing over video environment. In *Proceedings of the eleventh ACM international conference on Multimedia*. 100–101.
- [70] Ye Pan and Kenny Mitchell. 2020. PoseMMR: A Collaborative Mixed Reality Authoring Tool for Character Animation. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 759–760.
- [71] Danakorn Nincarean Eh Phon, Mohamad Bilal Ali, and Noor Dayana Abd Halim. 2014. Collaborative augmented reality in education: A review. In *2014 International Conference on Teaching and Learning in Computing and Engineering*. IEEE, 78–83.
- [72] Thammathip Piumsomboon, Arindam Day, Barrett Ens, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. Exploring enhancements for remote mixed reality collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*. 1–5.
- [73] Thammathip Piumsomboon, Gun A Lee, Jonathon D Hart, Barrett Ens, Robert W Lindeman, Bruce H Thomas, and Mark Billinghurst. 2018. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.
- [74] Thammathip Piumsomboon, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. CoVAR: a collaborative virtual and augmented reality system for remote collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies*. 1–2.
- [75] Thammathip Piumsomboon, Youngho Lee, Gun A Lee, Arindam Dey, and Mark Billinghurst. 2017. Empathic mixed reality: Sharing what you feel and interacting with what you see. In *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*. IEEE, 38–41.
- [76] VR Education Holdings PLC. 2020. EngageVR: Communicate, Teach, Learn. <https://engagevr.io/>
- [77] Erik Poppe, Ross Brown, Daniel Johnson, and Jan Recker. 2012. Preliminary evaluation of an augmented reality collaborative process modelling system. In *2012 International Conference on Cyberworlds*. IEEE, 77–84.
- [78] Erik Poppe, Ross A Brown, Jan C Recker, and Daniel M Johnson. 2011. A prototype augmented reality collaborative process modelling tool. (2011).
- [79] Lauriane Pouliquen-Lardy, Isabelle Milleville-Pennel, Francois Guillaume, and Franck Mars. 2016. Remote collaboration in virtual reality: asymmetrical effects of task distribution on spatial processing and mental workload. *Virtual Reality* 20, 4 (2016), 213–220.
- [80] Irene Rae, Bilge Mutlu, and Leila Takayama. 2014. Bodies in motion: mobility, presence, and task awareness in telepresence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2153–2162.
- [81] Jason Rambach, Paul Lesur, Alain Pagani, and Didier Stricker. 2019. SlamCraft: Dense Planar RGB Monocular SLAM. In *2019 16th International Conference on Machine Vision Applications (MVA)*. IEEE, 1–6.
- [82] Jason Rambach, Alain Pagani, and Didier Stricker. 2017. [poster] augmented things: Enhancing ar applications leveraging the internet of things and universal 3d object tracking. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, 103–108.
- [83] Holger Regenbrecht, Michael Haller, Joerg Hauber, and Mark Billinghurst. 2006. Carpeno: interfacing remote collaborative virtual environments with table-top interaction. *Virtual Reality* 10, 2 (2006), 95–107.
- [84] Holger Regenbrecht, Tim Lum, Petra Kohler, Claudia Ott, M Wagner, Wilhelm Wilke, and Erich Mueller. 2004. Using augmented virtuality for remote collaboration. *Presence: Teleoperators & Virtual Environments* 13, 3 (2004), 338–354.
- [85] Holger T Regenbrecht, Michael Wagner, and Gregory Baratoff. 2002. Magicmeeting: A collaborative tangible augmented reality system. *Virtual Reality* 6, 3 (2002), 151–166.
- [86] Linden Research. 2020. Second Life. <https://secondlife.com/>
- [87] Rec Room. 2020. Rec Room. <https://recroom.com/>
- [88] Michael Rymaszewski, Wagner James Au, Mark Wallace, Catherine Winters, Cory Ondrejka, and Benjamin Batstone-Cunningham. 2007. *Second life: The official guide*. John Wiley & Sons.



- [89] Bektur Ryskeldiev. 2018. Spatial Social Media: Towards Collaborative Mixed Reality Telepresence “On The Go”. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–4.
- [90] Bektur Ryskeldiev, Michael Cohen, and Jens Herder. 2018. Streamspace: Pervasive mixed reality telepresence for remote collaboration on mobile devices. *Journal of Information Processing* 26 (2018), 177–185.
- [91] Kyung Sakong and Tek-jin Nam. 2006. Supporting telepresence by visual and physical cues in distributed 3D collaborative design environments. In *CHI’06 Extended Abstracts on Human Factors in Computing Systems*. 1283–1288.
- [92] Prasanth Sasikumara, Lei Gao, Huidong Bai, and Mark Billinghurst. 2019. Wearable RemoteFusion: A Mixed Reality Remote Collaboration System with Local Eye Gaze and Remote Hand Gesture Sharing. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 393–394.
- [93] Alexander Schäfer, Gerd Reis, and Didier Stricker. 2019. Towards Collaborative Photorealistic VR Meeting Rooms. In *Proceedings of Mensch und Computer 2019*. 599–603.
- [94] Robin Schlünsen, Oscar Ariza, and Frank Steinicke. 2019. A VR Study on Freehand vs. Widgets for 3D Manipulation Tasks. In *Proceedings of Mensch und Computer 2019*. 223–233.
- [95] Yan Shen, SK Ong, and AYC Nee. 2006. A framework for multiple-view product representation using Augmented Reality. In *2006 International Conference on Cyberworlds*. IEEE, 157–164.
- [96] Yan Shen, Soh-Khim Ong, and Andrew YC Nee. 2008. Collaborative design in 3D space. In *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. 1–6.
- [97] Yan Shen, Soh-Khim Ong, and Andrew YC Nee. 2008. Product information visualization and augmentation in collaborative design. *Computer-Aided Design* 40, 9 (2008), 963–974.
- [98] Yan Shen, Soh Khim Ong, and Andrew YC Nee. 2010. Augmented reality for collaborative product design and development. *Design Studies* 31, 2 (2010), 118–145.
- [99] Ronald Sidharta, James Oliver, and Adrian Sannier. 2006. Augmented reality tangible interface for distributed design review. In *International Conference on Computer Graphics, Imaging and Visualisation (CGIV’06)*. IEEE, 464–470.
- [100] Doghead Simulations. 2020. Rumii. <https://www.dogheadsimulations.com/rumii>
- [101] SineSpace. 2020. Breakroom, The social hub for remote teams. <http://sine.space/breakroom>
- [102] Rajinder S Sodhi, Brett R Jones, David Forsyth, Brian P Bailey, and Giuliano Maciocci. 2013. BeThere: 3D mobile collaboration with spatial input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 179–188.
- [103] Maximilian Speicher, Jingchen Cao, Ao Yu, Haihua Zhang, and Michael Nebeling. 2018. 360anywhere: Mobile ad-hoc collaboration in any environment using 360 video and augmented reality. *Proceedings of the ACM on Human-Computer Interaction* 2, EICS (2018), 1–20.
- [104] Hongling Sun, Yue Liu, Zhenliang Zhang, Xiaoxu Liu, and Yongtian Wang. 2018. Employing different viewpoints for remote guidance in a collaborative augmented environment. In *Proceedings of the sixth international symposium of Chinese CHI*. 64–70.
- [105] Hongling Sun, Zhenliang Zhang, Yue Liu, and Henry BL Duh. 2016. OptoBridge: assisting skill acquisition in the remote experimental collaboration. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. 195–199.
- [106] Vive Sync. 2020. Vive Sync. <https://blog.vive.com/us/2018/11/09/vive-sync-announced-mainstream-enterprise-vr-collaboration/>
- [107] Spatial Systems. 2020. Spatial. <https://spatial.io/>
- [108] Matthew Tait and Mark Billinghurst. 2015. The effect of view independence in a collaborative AR system. *Computer Supported Cooperative Work (CSCW)* 24, 6 (2015), 563–589.
- [109] Zhipeng Tan, Yuning Hu, and Kun Xu. 2017. Virtual reality based immersive telepresence system for remote conversation and collaboration. In *International Workshop on Next Generation Computer Animation Techniques*. Springer, 234–247.
- [110] TechViz. 2020. TechViz VR. <https://techviz.net/>
- [111] Theophilus Teo, Gun A. Lee, Mark Billinghurst, and Matt Adcock. 2019. 360Drops: Mixed Reality Remote Collaboration using 360 Panoramas within the 3D Scene. In *SIGGRAPH Asia 2019 Emerging Technologies*. 1–2.
- [112] Theophilus Teo, Ashkan F. Hayati, Gun A. Lee, Mark Billinghurst, and Matt Adcock. 2019. A Technique for Mixed Reality Remote Collaboration using 360 Panoramas in 3D Reconstructed Scenes. In *25th ACM Symposium on Virtual Reality Software and Technology*. 1–11.
- [113] Theophilus Teo, Louise Lawrence, Gun A Lee, Mark Billinghurst, and Matt Adcock. 2019. Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [114] Theophilus Teo, Gun A Lee, Mark Billinghurst, and Matt Adcock. 2019. Merging Live and Static 360 Panoramas Inside a 3D Scene for Mixed Reality Remote Collaboration. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 22–25.

- [115] TheWild. 2020. TheWild: Immersive Collaboration for Teams. <https://thewild.com/>
- [116] Sebastian Utzig, Robert Kaps, Syed Muhammad Azeem, and Andreas Gerndt. 2019. Augmented reality for remote collaboration in aircraft maintenance tasks. In *2019 IEEE Aerospace Conference*. IEEE, 1–10.
- [117] VirBELA. 2020. VirBELA. <https://www.virbela.com/>
- [118] Arch Virtual. 2020. Acadicus. <https://acadicus.com/>
- [119] VR-ON. 2020. StageVR. <https://vr-on.com/>
- [120] vTime. 2020. vTime XR. <https://vtime.net/>
- [121] Peng Wang, Shusheng Zhang, Xiaoliang Bai, Mark Billingham, Weiping He, Mengmeng Sun, Yongxing Chen, Hao Lv, and Hongyu Ji. 2019. 2.5 DHANDS: a gesture-based MR remote collaborative platform. *The International Journal of Advanced Manufacturing Technology* 102, 5-8 (2019), 1339–1353.
- [122] Peng Wang, Shusheng Zhang, Xiaoliang Bai, Mark Billingham, Weiping He, Shuxia Wang, Xiaokun Zhang, Jiaxiang Du, and Yongxing Chen. 2019. Head Pointer or Eye Gaze: Which Helps More in MR Remote Collaboration?. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 1219–1220.
- [123] Peng Wang, Shusheng Zhang, Mark Billingham, Xiaoliang Bai, Weiping He, Shuxia Wang, Mengmeng Sun, and Xu Zhang. 2019. A comprehensive survey of AR/MR-based co-design in manufacturing. *Engineering with Computers* (2019), 1–24.
- [124] Xiangyu Wang. 2007. Exploring an innovative collaborative design space through mixed reality boundaries. In *2007 11th International Conference on Computer Supported Cooperative Work in Design*. IEEE, 264–269.
- [125] Xiangyu Wang and Rui Chen. 2009. An experimental study on collaborative effectiveness of augmented reality potentials in urban design. *CoDesign* 5, 4 (2009), 229–244.
- [126] Xiangyu Wang and Phillip S Dunston. 2006. Potential of augmented reality as an assistant viewer for computer-aided drawing. *Journal of computing in civil engineering* 20, 6 (2006), 437–441.
- [127] Xiangyu Wang and Phillip S Dunston. 2008. User perspectives on mixed reality tabletop visualization for face-to-face collaborative design review. *Automation in construction* 17, 4 (2008), 399–412.
- [128] Xiangyu Wang, Peter ED Love, Mi Jeong Kim, and Wei Wang. 2014. Mutual awareness in collaborative design: An Augmented Reality integrated telepresence system. *Computers in Industry* 65, 2 (2014), 314–324.
- [129] Wave. 2020. WaveXR. <https://wavexr.com/>
- [130] Tim Weissker, Pauline Bimberg, and Bernd Froehlich. 2020. Getting There Together: Group Navigation in Distributed Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics* 26, 5 (2020), 1860–1870.
- [131] WorldViz. 2020. WorldViz: Where virtual becomes reality. <https://www.worldviz.com/>
- [132] Boram Yoon, Hyung-il Kim, Gun A Lee, Mark Billingham, and Woontack Woo. 2019. The Effect of Avatar Appearance on Social Presence in an Augmented Reality Remote Collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 547–556.
- [133] Nadia Zenati, S Benbelkacem, M Belhocine, and A Bellarbi. 2013. A new AR interaction for collaborative E-maintenance system. *IFAC Proceedings Volumes* 46, 9 (2013), 619–624.
- [134] Nadia Zenati, Mahfoud Hamidia, Abdelkader Bellarbi, and Samir Benbelkacem. 2015. E-maintenance for photovoltaic power system in Algeria. In *2015 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2594–2599.
- [135] Nadia Zenati-Henda, Abdelkader Bellarbi, Samir Benbelkacem, and Mahmoud Belhocine. 2014. Augmented reality system based on hand gestures for remote maintenance. In *2014 International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE, 5–8.
- [136] Jakob Zillner, Erick Mendez, and Daniel Wagner. 2018. Augmented reality remote collaboration with dense reconstruction. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 38–39.