Review

A Review on Methods and Systems for Remote Collaboration

Romina Druta [1,*], Cristian Druta [2], Paul Negirla [1] and Ioan Silea [1,*]

[1] Automation and Applied Informatics Department, Faculty of Automation and Computer Science, Politehnica University Timisoara, 300006 Timisoara, Romania; paul.negirla@upt.ro

[2] Applied Electronics Department, Faculty of Electronics, Telecommunications and Information Technologies, Politehnica University, 300006 Timisoara, Romania; cristianfilip.druta@student.upt.ro

[*] Correspondence: romina.druta@student.upt.ro (R.D.); ioan.silea@upt.ro (I.S.)

Abstract: Due to the appearance of COVID-19 in 2019, person-to-person interactions were drastically reduced. The impact of these restrictions on the economic environment was significant. For example, technical assistance for commissioning or adjusting the parameters of some complex machines/installations had to be postponed. Economic operators became interested in the possibility of remote collaboration, depending on the manufactured products and the performance of the production lines that they owned. This bibliographic research was undertaken to address these needs. The purpose of this review was to analyze the current solutions, approaches, and technologies that workers and specialists can implement to obtain a reliable remote collaboration system. This survey focuses on techniques, devices, and tools that are being used in different contexts to provide remote guidance. We present communication cues and methods being employed, the implemented technological support, and the areas that benefit from remote collaboration. We hope that our effort will be useful to those who develop such systems and people who want to learn about the existence of collaborative solutions, and that it will increase awareness about the applications and the importance of the domain. We are convinced that, with the development of communication systems, the advancement of remote support systems will be a goal for many economic operators.

Keywords: remote assistance; smart device; video camera; mixed reality; virtual reality; augmented reality; communication cues; 360-degree panorama; annotation; object manipulation

1. Introduction

When talking about remote assistance, different situations requiring remote collaboration might come to mind. In some cases, work assignments, problems, or other issues might be easy to clarify during a phone call, but other tasks may necessitate the use of images or video communication so that an expert can understand each detail of a problem that an employee may be having. However, even a video call may not be sufficient for some tasks, e.g., if the problem is critical and highly complex.

There are multiple situations that may interrupt the expected continuation of a process, and in order to meet deadlines, they must be fixed immediately. The whole context is more complicated now due to the restrictions that arose from the COVID-19 pandemic. The reality is that there are not enough experts, and they are sometimes servicing or supervising multiple production sites. Thus, efficient collaboration systems for instant access are in high demand. Technical advancements and miniaturization enable real-time remote support, facilitating interactions and explanations for teams with workers located in different corners of the world. Infrastructure for remote collaboration enables fast problem identification and problem-solving, prompt intervention due to virtual presence, accessibility, and cost-effectiveness.

Remote collaboration is an emergent domain, and the tendency is for it to be democratized at a large scale due to technological advances and lowering costs.

Moreover, we must mention that privacy concerns, security threats, and standardization gaps are major concerns in this field but will not be detailed as they are not considered the subject of this paper.

As a general rule, industrial processes such as manufacturing, assembly, packing, etc., are in a race against the clock. To sustain the process and the infrastructure, facilities and installations need to function without error almost 24 h a day, 7 days a week. Given these time constraints, every outage must be immediately solved. Occasionally, detailed information is vital for a remote expert to comprehend the situation and guide a less experienced person accordingly. For example, during a remote assistance session, there is a fundamental need to identify specific targets, boundaries, devices, and parts. It is crucial to be able to discuss the defects, take corrective actions to remove the causes of the malfunction, and restore the system to a perfectly functioning state. Performing the correct actions is important as human error, accidents, and outages can damage devices or chains of systems. Overall, a holistic view of the system is needed so that, while corrective actions are taken to address specific problems, we keep in mind the whole picture of the process and are not limited to only the observed symptom or the small scope of the issue.

These remote assistance situations can be found in the medical field, such as surgery [1–9] and emergency management [10], in the education sector [11,12], mining [13], food industries [14,15], manufacturing industries [16–21], and crime scene investigations [22].

For the last 20 years, researchers have tried to find new ways to allow a remote (i.e., located in a different location) user/expert to help an on-site user by sharing the local user's view in real-time, thus improving the workspace awareness for the expert.

In early studies, different setups involving head-mounted cameras, head-mounted displays, video cameras, and pointing devices were developed to facilitate communication between two remote users. Lately, researchers have been focused on systems that rely on smart glasses and head-mounted devices based on augmented reality and virtual reality. Furthermore, there is interest from large companies around the world in developing devices that support real-time remote assistance and remote training for front-line workers [17–19,21]. Figure 1 shows the main devices used in remote assistance in the last few years, and depicts a collaborative scenario for connecting an on-site worker with a remote expert. The minimal equipment needed is either a PC with conferencing capability and remote control equipped with an LCD/console or a smartphone. A robotic arm on the worksite can be controlled remotely to point or move objects and ease the worker's tasks. The head-mounted display, the web camera, and the projector are useful for presenting the visual information used as the basis for giving specific guidance.
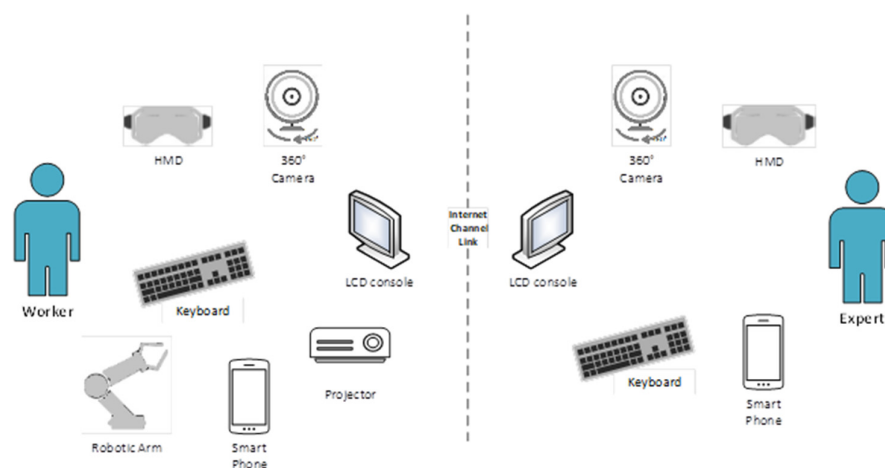

Figure 1. Worker side and remote expert side with devices enabling remote collaboration.

To the best of our knowledge, there has not been any comprehensive review of remote assistive technology until now. The motivation for conducting this review was the need to have an overview of the current technologies employed. There was no such type of review when we researched in the IEEE Xplore, Springer Link, and ACM Digital Library online databases. It is of great interest for researchers to have a clear view of the main methods and systems that would allow them to identify the position of their research in the field.

Therefore, this paper reviews the current methods and systems employed in remote assistance and collaboration to gain clear guidelines for future research.

We review state-of-the-art remote assistance technologies, devices used, visual display technologies, and communication cues.

2. Methodology

Our research methodology was organized in three phases: planning the review, conducting the review, and documenting the review.

The first step was to define the scope of the review by identifying the answers to the following research questions:

- "What kinds of devices are used in remote collaboration?"
- "What are the industrial domains making use of remote collaboration infrastructure?"
- "What are the technologies used in displaying the information?"
- "What are the types of communication cues used in remote collaboration?"
- "What can be improved to offer a better user experience?"

The second step was to conduct the review by selecting the existing studies in the field, extracting the required content, and synthesizing the data for the last stage of elaborating the review.

A total of 230 articles were selected by querying two databases, IEEE [23] and ACM [24], using the following keywords: "remote collaboration", "remote collaboration system", "remote video collaboration", "remote assistance", and "remote guidance". We chose these two databases because the number of articles that resulted from the queries was significant, and we considered that the content was relevant for our research.

Figure 2 shows the dynamics of the "remote collaboration" subject in academia and industry. Figure 2 depicts the distribution graph of the relevant articles sorted by database and year. Starting in 2010, a steadily increasing trend can be observed, almost tripling from 11 works (3 in IEEE and 8 in ACM) in 2010 to 31 works (6 in IEEE and 25 in ACM) in 2020. For the keywords "remote collaboration" and "remote intervention", Google Trends records data starting in 2004 [25]. We observed some overlap during 2004, when the interest was quite high, until early 2005. This interest can be explained by the first preoccupations in remote collaboration using augmented reality. As with each new technology that appears, the interest was high in the first year; afterwards, it decreased but had occasional spikes between 2005 and 2014. Interestingly, in March 2020, a sudden interest could be observed, which coincided with the beginning of the COVID-19 pandemic.
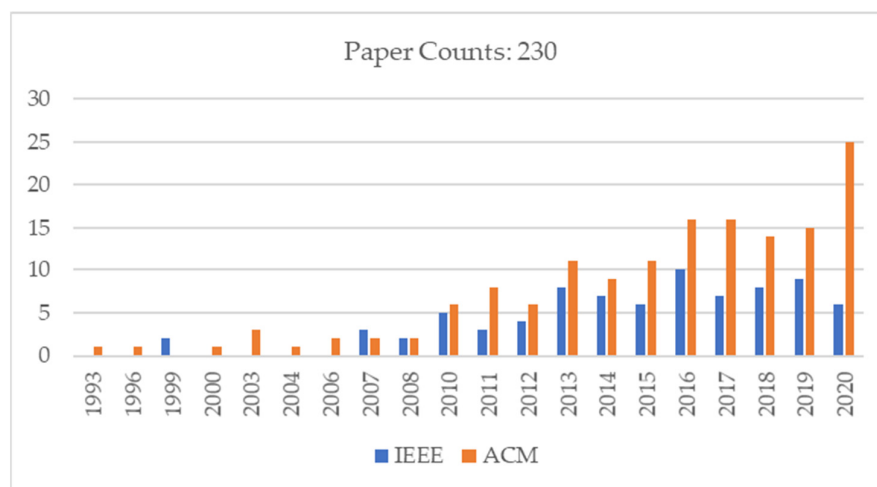
Figure 2. Distribution of publications for the subject "remote collaboration" during 1993–2020.

From the total of 230 articles found in the two indexed databases, we selected the most relevant ones, excluding the papers focusing on videoconferencing, thus retaining 118 papers pertinent to the subject of remote support. After extracting the most valuable information from the selected articles, we identified three main directions for this review. Accordingly, the content of this survey is arranged on three axes, which correspond to the topics of technologies in the remote assistance sector. The first motif is represented by the electronic devices that serve the remote assistance task. Secondly, we deepened the display technologies embodied in the devices employed, and, finally, we focused on the communication cues as means used by the expert for guiding the remote user to identify objects and follow instructions. We depict in Figure 3 the main topics of this study, which will be addressed in the following sections: existing electronic devices, display technologies, and communication cues.
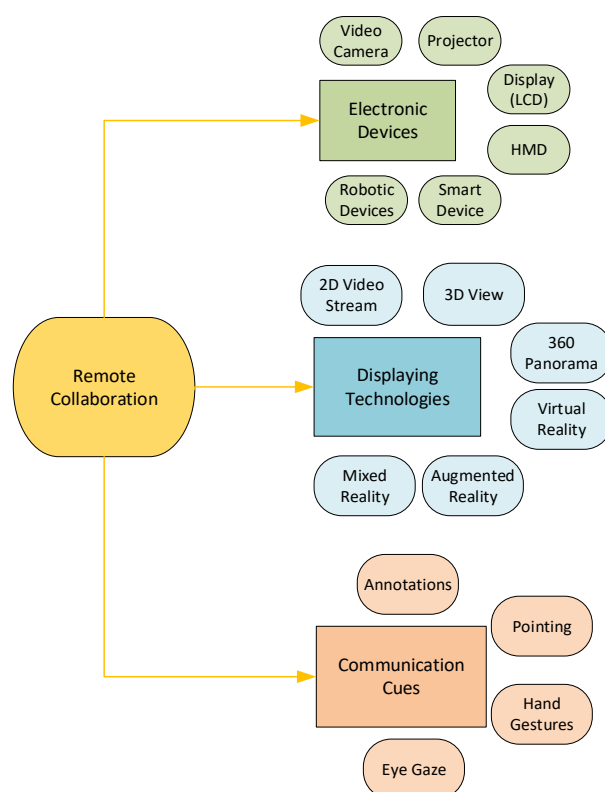
Figure 3. Plan of the review distributed across the three focus areas—electronic devices, display technologies, and communication cues—with the generalization of each topic into sub-topics.

For each of the three aforementioned topics, the information was systematized in a table that identifies the works in a specific scientific field and the different methods/equipment used. Thus, future researchers will be able to more quickly locate the scientific works and the methods/equipment used by different authors for a certain field/task/problem. Readers will be able to focus only on the solutions that apply to their activity, and they will be able to identify the advantages or disadvantages of the various existing/applied solutions from the comparison done in this review.

3. Electronic Devices Used in Remote Assistance

When developing a remote collaboration system, particular devices with different form factors are adapted for various purposes and scenarios. For example, video cameras can be head-mounted or worn on the shoulder; information can be displayed on a PC screen, projected, or seen through a head-mounted device (HMD).

The system proposed by Villaruel et al. [1] was intended for remote surgery. A 2 DOF (degree of freedom) robotic arm was controlled distantly by electromyographic signals captured from the remote user's muscles, corresponding to the human arm's natural movements.

Another system for remote surgery is presented in [2]; the doctor performing the surgery uses an HMD and controls a surgical robot remotely. The article's title is "The rise of robots in surgical environments during COVID-19", which underlines the favorable context for the proliferation of remote systems during the COVID-19 pandemic.

A more complex system used in remote surgery that employs a robotic arm is presented by Suthakorn et al. in [4]. The system has an expert station that allows the surgeon to control the surgical robots located at the remote intervention site through a 3D force-feedback haptic robot.

M. Bauer [26] was among the first researchers who proposed a system to be used by two remote users for educational purposes, the aim being to set up a wired circuit. The prototype developed for the worker's side was composed of a video camera attached to a head-mounted device (HMD). The camera sensor was pointing away from the user in the approximate direction of the local user's gaze and tried to capture images of the task area. On the remote side, the expert analyzed the images and instructed the user on what steps to follow via a telepointer.

Susan R. Fussell's [27] study showed that in a remote bike repairing context, in addition to a head-mounted camera worn by the repairer, a scene camera provided valuable visual information for remote collaboration on physical tasks and augmented the feeling of co-presence for the remote helper.

The wearable active camera laser (WACL) described in [28] incepted the idea of placing the camera on the shoulder of the local user. A laser pointer was attached to the video camera, and the remote user controlled the laser beam intended for spatial guidance. The expert and the worker had headsets with microphones used to convey instructions and explanations from the working site. The system was tested in the context of an assembly task (Lego assembly). The results from [28] showed that, compared with a head-mounted camera, the WACL was more comfortable and caused less fatigue, although, if the wearer moved the laser pointer slightly, the pointer position changed; hence, guidance became difficult.

Machino [29,30] proposed a 2 DOF robot on wheels, situated on a worksite and controlled by a remote expert. A video camera was fixed to the top of the robot, allowing it to capture the local environment, while a projector was used to project the expert's instructions to the remote worker. The solution was targeted to the manufacturing industry for tasks involving maintenance operations.

M. Adcock and C. Gunn [31] presented a remote guidance system that enabled a remote expert to instruct a mobile worker using sticky annotations. The worker wore a helmet equipped with a small camera and a laser pico projector placed on top of the helmet. These devices were connected to a laptop in the local user's backpack. The remote helper received the video stream on a tablet that also communicated with the worker's PC. The tablet allowed the expert to sketch the instructions, which were transmitted, with the help of the pico projector, as a 3D line into the local user's environment. This system was aimed to be deployed in the domains of industrial manufacturing for maintenance and healthcare for medical consultations.

Starting in 2010, we noticed that smart devices (mobile phones or tablets) were designed to facilitate remote visualization and collaboration between workers in factories and experts in different locations [14,15].

The prototype proposed by Gauglitz in [32] is a mobile remote collaboration system used to assist a user in operating a mock-up airplane cockpit. The pilot was guided by the instructions of a remote helper. The local user was equipped with a tablet device with an integrated camera for sending environmental information remotely, and the display showed the virtual annotations placed by the remote expert. The remote user was presented with a view of the local user's environment and, besides controlling the annotations, the helper could also "freeze" his viewpoint at any time. Later, Gauglitz et al. [33] presented an improved solution for mobile remote collaboration. The solution enabled the helper to have an independent view of the local user's environment by controlling the remote video camera and the augmented reality (AR) annotations displayed on the local user's smart device (tablet). It was tested by a user repairing a car following the instructions of a remote helper.

Teleadvisor [34], a novel design used to solve complex wiring tasks, comprises two devices, the Teleadvisor component and the controller display. A worker controls the position of the Teleadvisor, represented by a robotic arm with a camera and a pico projector mounted on the top. Teleadvisor's camera captures the visual information, and the data are transmitted to an off-site controller operated by an expert. With the help of

the projector, the information sent from the controller back to the Teleadvisor is displayed as an overlay emphasizing objects in the worker's environment. Thanks to the mobile robotic arm, the helper can move the camera's viewpoint closer to the object and zoom in and out.

All the aforementioned solutions [26–34] require a fixed setup for the cameras, either on the worker's or the helper's side. In contrast, HandsInAir [35] includes a new feature that allows mobility for both the remote helper and the worker. This novel approach is highly relevant for factories where workers need to walk around and inspect different machines. The solution requires that both users are wearing the same equipment, consisting of a camera mounted on a helmet and a custom-built near-eye display beneath the brim. The worker's camera is used to capture the worker's actions, and the remote user's camera captures the hand gestures of the helper, providing instructions for the local user.

The MobileHelper presented in [36] proposes a similar approach as [35]. The worker has a helmet equipped with a video camera and a near-eye display, but the remote helper uses a tablet. The video transmitted from the worker's camera is displayed on the helper's device. The tablet's video camera captures the hand gestures of the helper and combines them with the video data, and the result is displayed on the helper's tablet screen and the worker's near-eye display.

The study in [37] focused on how AR can be used to enhance remote space collaboration. The system used a Kinect depth sensor to capture a 3D model of the local user's surroundings, on which a remote user could overlay annotations. These labels appeared projected onto a screen with the help of a laser projector, which in turn could be manipulated by the remote expert. This solution also allowed the possibility for the off-site expert to manipulate the scene independently of the view of the local user.

Domova [38] presented a system conceived for solving physical tasks in the manufacturing industry, allowing both the on-site worker and the off-site helper to annotate video feeds. The worker carried a mobile phone with video streaming capabilities to capture the details of the local environment, and the scene information was displayed on the remote user's PC. Both users could take snapshots, and the stream could be frozen on either side when one of the users wanted to point at a particular scene at a chosen moment.

PopArm [39] is a solution based on a robotic arm that seems to pop out from a video streamed between two remote users. The robot arm is synchronized with the remote instructor's arm movements, and it moves and rotates on the local user's display, enabling the off-site user to point to and touch remote objects in the local user's environment. PopArm is useful when there is a need to point to specific objects remotely.

RemotIO [40] consists of two head-mounted cameras, one for the helper and one for the novice user. The local user shares his view; the video is then used for immersing the expert into the worker's field of view. The expert's hands are tracked with a depth sensor and superimposed in real time onto the worker's environment. The local user's hands and the superimposed expert's hands can be seen by both users. RemotIO aims to be used in the industry for remote maintenance and repair.

ExpertOnWheels [41] is a mobile telepresence robot designed to support collaboration between a field worker and a remote expert in the manufacturing industry. The architecture of the collaboration system is showcased in Figure 4. The mobile robot comprises a video camera for capturing the environment's details, a projector used for annotations, a speaker, a microphone for sound, and wheels for movement. A computer is the processing brain of the ExpertOnWheels (worker's side), which receives commands via the internet from the remote expert's PC controls. The local worker moves the robot only where intervention is needed.
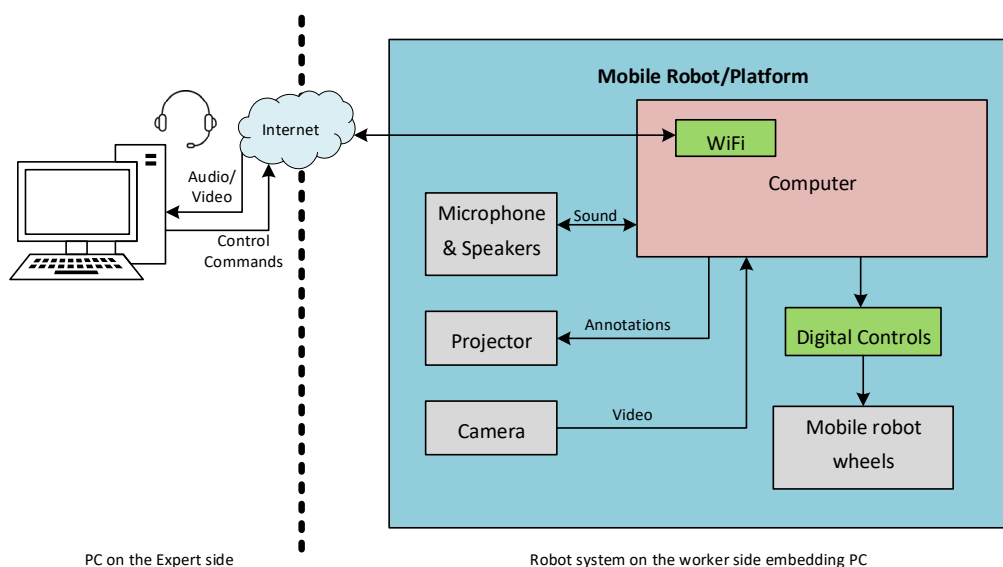
Figure 4. System architecture diagram for ExpertOnWheels [41].

To determine the task efficiency when collaborating remotely, Johnson et al. [42] conducted a comparative study for the use of handheld devices and head-mounted devices (HMD). The conclusions showed that, for static tasks, handheld devices were more suitable, but for dynamic tasks, head-mounted devices helped the user to finish a task faster. In another article, Johnson explored how the field of view impacts collaboration [43]. In this study, a robot controlled by a remote participant captured the distant environment at different angles. The results indicated that using a wide angle and a panoramic view contributes to quicker task completion.

The JackInHead system [44] tries to overcome the field of view limitation of the remote user by using headgear with multiple cameras to capture an omnidirectional video. The local user wears the headgear, while the remote user wears a head-mounted display with head tracking functionality.

Tait and Billinghurst [45] investigated how the independent view helps users to finish tasks faster when using an AR interface. The prototype allowed a remote user to navigate independently in the local user's space using a 3D scanned model of the local user's environment. The local user wore a head-mounted display that helped to render the annotations of the remote user. A four-video-camera system was used to track the objects from the local environment and the head movements of the host.

Smart Phone/Pad and Robot for Tele-operation and Tele-presence (SPRinT) [46] is another remote collaboration system using a robot located at a local user's site. The robot is equipped with a video camera and a projector to display the instructions of the remote user. The local user can adjust the robot and the projector position while the expert controls it using a mobile phone.

ADAMAAS [47] aims to determine the local user's actions in order to trigger instructions that are displayed as visual images or text on the HMD worn by the respective user. The system is composed of a head-mounted display with AR capabilities and an eye-tracking module.

In order to increase the situational awareness of the remote user, S. Kratz and F. Rabelo Ferriera [48] focused their research on improving the user's view using a mobile telepresence robot situated at the worker's site. The study's main interest was to see how quickly the job was completed when the worker was using an HMD with head-tracking and mono camera systems, while the helper was using a fixed 2D monitor to see the worker's space. The video feed, captured at the local site by a camera mounted on the robot pan/tilt servo system, was seen by the remote user on his HMD. Using a head-

tracking module, the helper could control the robot and the camera's orientation so that he could see the remote environment from various angles.

Most of the previously described prototypes concentrate on the remote user's awareness rather than the local user's reactions. In contrast, Empathy Glasses [49,50], proposed by Lee et al., focus on acquiring the local user's heart rate and facial expression as well as his gaze and viewpoint. To accomplish this, the system combines a see-through display with a head-mounted camera, an eye tracker, a facial expression tracker, and a heart rate monitor. The helper sees all the information acquired from those modules on a remote desktop interface.

Traditional remote collaboration solutions have examined what a single user sees in his local space, but JackInSpace [51] utilizes a new approach that allows a remote user to switch between the view of different local users (body users). In the prototype, the local user's side includes a head-mounted fisheye camera and depth sensors that capture the data and recreate a virtual 3D environment for the ghost (remote user). The remote user can enter the first local user's view to analyze the environment and, afterwards, can switch to another local user's view to see the scene from a different angle. The ghost wears 3D glasses to see the virtual perspective of the body user, displayed on three screens with the help of three projectors, an HDMI splitter, and a motion capture system.

MirrorTablet [52] is a low-cost system used for capturing hand gestures and is proposed for situations where both the helper and the worker are equipped with a mobile device (tablet). The hardware setup on the helper's side is composed of a mobile tablet on top of which is positioned a mirror within the field of view of the tablet's video camera. The mirror reflects the tablet's screen and the image of the hand gestures to make them visible to the camera. The video data captured from the worker are transmitted to the helper, and the instructions from the helper are overlaid on the worker's screen.

Unlike traditional remote collaboration systems, where the communication is done face to face, Gutsy-Avatar [53] allows the remote user to have the same view as the local one in an unusual way. This is performed using an electronically controlled t-shirt. The system architecture is presented in Figure 5. A local user wears a t-shirt embedded with a smart device/tablet with an attached camera. The device captures the local user's surroundings and displays them to the remote user. Behind the tablet, there are four servo motors commanded by a microcontroller, which receives commands from the expert's PC via the user's tablet. For the remote user, a PC with a video camera is needed to capture the user's face and send instructions related to the direction of the local camera view. The remote user changes the camera's view by sending commands to the server.
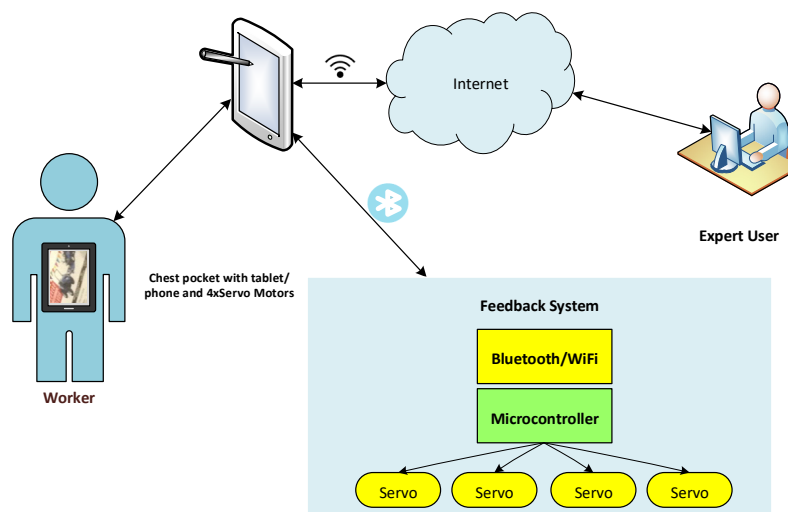

Figure 5. Gutsy-Avatar [53] system diagram.

Microsoft Remote Assist [17–19] is based on a HoloLens app that enables the remote collaborator to see everything that the local user is seeing (including holograms and the real environment) and to add annotations that can be seen by the local user. The host wears a Microsoft HoloLens device equipped with a video camera and depth sensors capable of tracking hand gestures and eye and head movements.

Remote Manipulator (ReMa) [54] focuses on reproducing the local object's manipulation of the remote site using a proxy object. At the tracking site, the object's manipulations are captured by infrared cameras and are transmitted to the manipulator's site, where they are displayed on a similar object, called a proxy object. The proxy object is automatically oriented by a robotic arm to reflect the position from the tracking site.

SharedSphere, proposed by Lee et al. [55], uses mixed reality (MR) to add gestures and cues to a 360-degree live panorama video. The local user wears a see-through AR HMD with an attached 360-degree video camera. The guest user wears a virtual reality (VR) HMD to watch the live panorama scene. A head tracking sensor is mounted on the host's side so that the remote user can have an independent view of the local user's environment even if the local user moves his head. The system uses view frames to indicate where each user is looking.

Unver et al. [56] addressed the use of multi-cameras for remote collaboration. The study compared the efficiency of using handheld devices over the hands-free experience in the context of using multiple video cameras. The system is composed of three types of cameras focused on the room (space), task, and the collaborator's face. Cameras could be switched manually or automatically depending upon the activity in the room. The study showed that task performance was the same when using handheld devices compared with when the user had his hands free, but users preferred the flexibility of the hands-free setup.

In the context of an industrial environment, S. Kesavan [16] proposed a video collaboration platform for remote interaction based on mobile devices and web cameras. The system has three main components: a dashboard receiver, a cloud service, and a plant operator. If a worker needs assistance in configuring industrial equipment, the machine's information is loaded onto a dashboard receiver, where it can be accessed by a remote expert. A cloud service hosts the APIs, which the collaborators can access to connect to the dashboard and operate. The system allows the operator to connect with the expert using a tablet device with a camera and to point the camera at the malfunctioning equipment. The collaborators can exchange information through video and audio channels and text messages through the chat functionality.

Kwon et al. [57] defined a fully asymmetric remote collaboration system where the remote user identified the local user's problem and had full control of the worker's environment while the worker had limited responsibilities. On the local user's side, a wheeled robot controlled remotely by a handheld device was equipped with a projector, a high-resolution camera, and a 360-degree camera. The remote user wore an HMD device that displayed information about the local user's surroundings captured by the robot. The remote user was capable of zooming in and out, pointing to objects, or adding annotations that could be projected onto the local user's environment.

The system proposed by Teo et al. [58] is composed of an AR HMD device with a 360-degree camera on top, on the local user's side, and a VR HMD on the remote user's side. The local user captures 360-degree panoramas of the local environment and sends them to the remote user, where they are displayed as a prebuilt 3D scene. Verbal communication is accomplished through the speakers of the HMD, and for non-verbal communication, the remote user can use the VR controller for pointing and instructing the local user.

OmniGlobe [59] uses a VR HMD device on the worker's side and, on the other side, a spherical display with a 360-degree camera, which is located on a rotating platform, allowing the specialist to move around. The system has a 360-degree first-person mode that allows the expert to see the VR environment from a panoramic view, and a third-

person mode makes it possible to see the environment from a higher elevation. Communication is enhanced by the system's ability for the users to share body gestures, gaze cues, and facial expressions.

Mohr et al. [60] proposed an MR collaboration system using mobile devices that uses light fields for orientation. The system records the local space and sends images to the remote user, who annotates them and sends them back to be visualized in AR by the local user.

The following table (Table 1) presents the repartition of the aforementioned articles and the system components used for the proposed solutions. Interested readers can easily and quickly identify the works that address certain areas/tasks and the methods/equipment on which the solution of remote assistance/collaboration is based.

Table 1. Devices and components for remote collaboration systems.

Application (Task Type)	Devices for Capturing and Visualization						
	Video Camera	HMC [1]	HMD	Smart Device [2]	Robotic Device	Projector	Display
Surgery	[1,2,4,7,8]		[2,7,8]		[1,2,4]		[1,2,4]
Crime Investigation	[26]		[26]				[26]
Building Wired Circuit		[27]	[27]				[27]
Lego Assembly	[28,35]	[35]	[28,42]	[42]			[28,35,42]
Education	[29]					[29]	[29]
Physical Task—Industry	[30]				[30]	[30]	[30]
Medical Consultation		[31]					[31]
Aviation				[32]			[32]
Car Maintenance	[33]			[33]			
Wiring tasks	[34]				[34]		[34]
Object Manipulation	[36,37,53]			[38]	[36,54]	[37]	[36–38,54]
Object Pointing	[39]				[39]		[39]
Maintenance—Industry	[41]	[40]		[60]		[41]	[40,41]
Object Identification	[55]	[54,58]	[43–45,54,58,59]	[46,55]	[43,46]	[45,46,59]	[45,46,54,55]
Cooking			[47]				
Assembly—Furniture	[48]		[48]		[48]		[48]
Assembly Task	[57]	[49–51]	[57]	[52]	[57]	[51,57]	[51]
Sport—Education				[53]			[53]
Manufacturing	[16]			[16]			

[1] Head-mounted camera; [2] Mobile phone/tablet.

4. Visual Display Technologies

In this section, we focus on methods that are used to share the visual space between two distant collaborators.

The most common method used in remote assistance and collaboration is a simple video camera with the ability to stream a 2D video, which is shown on a monitor at the helper's side. In order to give the remote collaborator the feeling of being collocated with the local user, the researchers also experimented with the use of 3D videos, 360-degree panoramas, virtual reality, augmented reality, and mixed reality. For each of the last three technologies, we will also provide a short definition within our discussion for a better understanding.

4.1. 2D View

A range of existing systems facilitate remote collaboration on physical tasks using 2D videos, as shown in the research articles [15–17,21–25].

Most remote collaboration systems use video cameras installed at the local user's (worker's) space (environment). The video captured by these cameras is shared in 2D on a display such as a PC, TV, or smartphone. However, this method does not allow the remote user to understand the spatial relationship between the objects; thus, communication is more difficult.

4.2. 3D View

Studies have been performed to obtain a 3D view that is presented to both operators. In one of these studies, G. Welsch [61] presented a prototype that uses an array of cameras to obtain a three-dimensional view for a remote environment.

OmniKinect [62] also tries to obtain a 3D view, but instead of using multiple cameras, it uses Kinects, which have depth sensors, to acquire real-time video.

The technology 3D Helping Hands [63] uses special cameras that are capable of 3D video capture in real time on both the worker's and the helper's side. The 3D videos acquired from the worker and the helper are then fused in real time to form a single, common workspace with an augmented view.

A new prototype that uses depth cameras at the worker's space [64] can capture in 3D the worker's position and environment. The cameras are connected to a fusion PC that renders the video stream with the helper's visual instructions.

Gauglitz [65] presented a new model of a remote collaboration system, where the local user has a smart device (tablet or smartphone) that captures the video information and sends it to the distant user. The remote user uses AR and adds annotations, which are overlaid on the source video and displayed on the local user's device.

Gao et al. [66] presented a mobile-based collaboration system in which the local user has a mobile phone equipped with a depth sensor in order to capture the local environment and display it in virtual reality (VR) as a 3D environment for the remote user. The system allows the expert to move independently in the virtual environment and to switch the view of the local user's smartphone to show the local changes in real time in the form of a 2D video.

4.3. Virtual Reality (VR)

Virtual reality is a visual technology that allows a person to immerse themselves completely in a digital world, independent of the physical environment, using special devices. The digital environment is computer-generated and contains objects and scenes that appear to be real. These environments can be used in training, games, or other live broadcast events. The user is equipped with a VR headset to access the applications that support VR. Popular VR devices include Oculus Rift [67], HTC Vive [68], and Oculus Quest [69].

Traditionally, immersive technologies have been mostly used to facilitate guidance during the exploration of an unfamiliar environment. From 1970 to 1990, VR devices were used for flight simulation, military training, and medical purposes. In the decade following 1990, these devices started to be commercialized for entertainment purposes. In recent years, VR devices have become more affordable. The combination of affordability and COVID-19 restrictions have fueled an enormous rise in their use [70]

CoVar [71] is a mixed reality system that supports collaboration between augmented reality and virtual reality users by sharing a 3D reconstructed environment. The VR user wears an HTC Vive Lighthouse device, which displays a reconstructed mesh. The AR user captures his real, local environment and shares it with a remote augmented virtuality user. Both collaborate on the tasks in the shared space. The system enhances the communication by using awareness cues (field of view and gaze cues) and an AV-Snap-to-AR interface, which enables the AV user to control his head orientation. The interaction between the users is improved by eye-gaze, head-gaze, and hand gestures.

Gao et al. [72] presented a prototype for remote guidance that uses VR headsets for both the host and the remote user. The workspaces of the worker and the helper are

captured using camera depth sensors, which results in sharing 3D environment data for the remote user, who can also track the worker's viewpoint. At the remote user's site, the sensors capture the helper's hands to indicate the instructions that have to be followed by the worker. Thus, in the shared virtual space, the worker sees the helper's hands overlaid on his own hands.

The study in [73] focused on allowing the helper to have an independent view and study objects from any direction in a shared environment by capturing and reconstructing a copy of the local worker's space in 3D. Before the collaboration starts, the local user moves around their space using the VR device to capture all the details of the local environment. The remote helper can see the scene in a VR environment.

Elvezio et al. [74] proposed a remote collaboration in AR and VR intended to allow the remote expert to create or manipulate virtual replicas of the real objects. The remote user gives instructions to a local user by manipulating, pointing, or annotating the virtual objects in his VR environment. The same annotations appear in the AR environment of the local user. Both users wear a head-worn display (HWD).

4.4. Augmented Reality (AR)

Augmented reality places digital objects or annotations onto the real world. Historically, augmented reality was used with HMD devices capable of displaying digital content overlaid on a user's view of the real world.

In 1957, augmented reality appeared in the form of the Sensorama [75], an invention that could deliver visuals, sounds, vibrations, and smells to a viewer. Nevertheless, the term AR started to be used in 1990, when several workers wearing HMDs were guided in assembling electrical wires in aircraft. Over the last decade, different applications and devices that employ AR have emerged; for example, the design tool ARToolkit appeared in 2009, Google Glass in 2013, and HoloLens in 2015 [75].

One system that implemented AR was the Télé-Assistance-Collaborative system developed by Bottecchia [76], in which the operator was equipped with a specific AR display device. Its design enabled it to capture a video flow of exactly what the carrier's eye saw (flow A) and a wide-angle video flow (flow B). On the other side, the expert had a catalogue of images and annotations that he could use by applying them on the video stream seen in real time by the operator.

J. Gu [12] presented another solution where AR was used with mobile devices to overlay objects onto the real environment.

In most collaboration systems where AR was utilized, head-mounted displays were employed. Schneider et al. [77] proposed an AR application that used edge computing. This solution consisted of three main components. The local user had a mobile AR device that captured images of the site and displayed the incoming video. These images were processed by an edge server, which sent them to the remote user and then back to the operator after overlaying the expert's annotations. The helper used a laptop/PC device that displayed the incoming video flux and a software program that allowed him to draw instructions on the images.

Zillner et al. [78] focused on the dense reconstruction of the local user's surroundings by generating a three-dimensional mesh automatically. The local user wore a pair of AR glasses equipped with a depth sensor used for scene reconstruction. The remote user could explore the scene independently, add annotations, and create 3D animations.

SceneCam [79,80] was a multi-camera remote collaboration system that allowed a remote user to have multiple views of the task space. The local user wore an AR HMD that collected information about the local space and ran an algorithm that determined which camera view captured his actions the best. Based on this selection, the system drew the remote user's attention to the optimal view. Another technique implemented by the system was the automatic camera view selection, which implies that, based on the algorithm, the system selected the optimal view for the remote user, making it their primary view.

4.5. Mixed Reality (MR)

Unlike VR, which implies a complete immersion experience that shuts out the physical world, mixed reality (MR) is a method that mixes the real world with the artificial world. Digital objects are superimposed, and images are overlaid onto the real world. MR systems enable distant collaborators to feel as though they are in the same space. The user needs a headset capable of MR (e.g., Microsoft HoloLens) in order to experience the involvement of virtual objects with the real world or physical view.

Yang and Peng in [81] presented a remote collaboration solution based on mixed reality. The real objects were virtualized using computing virtualization and displayed at the remote user's site. The local user could see how the remote user interacted with the virtual objects and could take the same actions on the real objects.

To increase the collaborators' shared view and reduce the ambiguity of deictic expressions, Muller et al. [82] proposed the use of shared virtual landmarks in a mixed reality environment.

Feick et al. [83] proposed a design for an MR system where both the novice and the expert wore head-mounted devices with video capturing capabilities. Based on the object of interest from the novice's environment, a proxy object is created in a virtual environment that can be seen by the expert in a split display together with the remote live video feed. The novice can also see the virtual proxy object and the expert's gestures.

Teo et al. [84,85] proposed an MR collaboration system that used a live 360-degree panorama in a 3D reconstructed scene. A remote user could choose the way in which he interacted with the local user's environment, either via a live 360-degree panorama video, using past static images captured from the 360-degree panorama, or using the 3D reconstructed scene. In [86], the authors compared the results obtained when allowing the user to switch between the 3D view and the 360-degree panoramic view of the local environments. The results showed that the users reported that their concentration was better when solving their task using the 3D mode and had a good understanding of the collaborator's focus in the 360-degree mode.

In addition to the capabilities of the system proposed by Teo et al. in [84], Gao et al. [87] allowed the local user to share a 2D first-person view of the local environment. The local surroundings were captured as a 3D scene and displayed on the remote VR HMD. The expert could analyze the virtual space from three different perspectives: the 2D first-person view, a 3D static view, and a 360-degree panoramic view. The results showed that the users preferred the 360-degree view, as it offered more control and independence compared to the 2D view.

4.6. 360-Degree Panorama View

Another display option for collaborating remotely is the 360-degree panorama view, where the expert user can see all the surroundings of the local user.

LiveSphere [88] is a remote collaboration solution that consists of a system with wearable camera headgear that provides 360-degree spherical images of the local user's surrounding environment. The headgear has six video cameras, and the video streams are fused into a spherical video that is transmitted to a remote user. This setup allows the off-site user to see the local user's view and also to have the freedom to look around via the on-site user's view.

JackInHead [89] allows a local user to send an omnidirectional video to a remote user, who has an independent view of the local user's space. The prototype involves the use of headgear with multiple cameras worn by the local user and an HMD worn by the remote user. The omnidirectional video resulting from the images captured by the headgear is sent to the remote user, where it is mapped in a spherical virtual space.

SharedSphere [90,91] is a system used in interactive collaborations and allows 360-degree panorama video live sharing between two users. Hand gestures are overlaid on the live video; thus, mixed reality is used to enhance the collaboration. The guest user

wears a VR HMD on which a high-resolution camera and a second 360-degree panorama camera are positioned. The video captured at the local site is sent to the host user, who wears a see-through HMD. The guest user can see the local user's hand gestures thanks to the video cameras, but the hands of the guest user need to be tracked with a hand-tracking sensor in order to be seen by the host user.

Kangas et al. [92] proposed a remote collaboration solution that enabled the remote expert to have a 360-degree view of the remote environment while instructing the local user with the help of a projector equipped with a close-up attached camera and pointer capabilities. This setup allowed the expert to analyze the local user's environment independently.

In contrast with other remote collaboration systems that offered a 360-degree view that allowed interaction between only two users, 360Anywhere [93] allows multiple users to interact remotely by annotating a 360-degree video and projecting it back to the local users using AR. The remote collaborator has a total view of the local environment without the need for the local user to move or adjust the cameras. Besides adding annotations, the user can rewind the video, chat with the remote user, calibrate the system, and track the collaborator's gaze.

On the Shoulder of the Giant [94] is a multi-scale mixed reality system sharing a 360-degree panoramic video that facilitates collaboration between a local VR user and a remote AR user. The study presents two modes of collaboration realized by the system: one shares a 3D reconstruction with the remote user, and the other shares a panoramic video at a different scale, using a 360-degree camera controlled by the remote user. Both users are allowed to switch between the two modes of collaboration. The use of the 360-degree camera was tested in different positions: head, back, hand, shoulder. The most preferred camera position was on the shoulder, which allowed the user to see the collaborator's face and environment.

Table 2 shows the repartition of the articles and the display viewing technologies used in remote collaboration in the application field. It can be seen that, for maintenance, object manipulation, or other physical tasks, the visual display methods are important for both remote collaborators to have a clear view of the environment or the instructions that have to be followed.

Similar to the previous section, in the following table, interested readers can easily and quickly identify the papers that address certain areas/tasks and the methods/equipment on which the solution of remote assistance/collaboration is based.

Table 2. Visual display methods used in remote collaboration.

Application (Task Type)	Visual Display Methods					
	2D View	3D View	VR	AR	MR	360-Degree Panorama View
Healthcare		[61]				
Object Reconstruction		[62]		[62]		
Assembly Task		[63]				
Physical Task	[64]			[64]		
Car Maintenance	[65]			[65]		
Object Manipulation	[66,79,80,87,92]	[66,79–81,83–87,94]	[66,83,87,94]	[66,79,80,83,87,93,94]	[66,81–87]	[84–87,93,94]
Lego Assembly		[72,73]	[72,73]		[72,73]	
Maintenance	[74]	[74,78]	[74,78]	[74,78]		
Maintenance—Industry	[77]			[77]		
Object Identification			[71]	[71]	[71,90,91]	[88–91]

Maintenance—Hydroelectric Power Plant	[92]	[92]

As can be seen in the table, in this context (visual display technologies), most of the researchers focused on systems that allow a helper to give instructions on tasks involving object manipulation.

5. Communication Methods and Cues

In face-to-face collaboration, for both collaborators involved in exchanging information, not only what the person is saying but also where they are looking and what their hands and body gestures are showing is important.

Fussel [95] showed that for remote collaboration, video communication is better than having only audio. For a helper–worker pair involved in a bicycle repair task, Fussel compared the task performance in the following scenarios: one where the communication was done using visual information, and another where the communication was audio-only. The conclusions showed that a shared visual space was essential for the collaborative task.

In addition, in order to establish common ground, users need to use communication cues such as speech and no-speech audio, gaze, facial expression, and hand and body gestures. Many of these cues can be captured using different video telecommunication systems, such as Skype, Google Hangouts, or Zoom. However, in some cases, the remote user needs to point to a specific object or draw specific instructions; here, pointing gestures and representational gestures (annotations) are important [96]. Kirk and Stanton demonstrated how remote gestures influence the structure of collaborative discourse and how their use can also influence the temporal nature of the grounding process [97].

In [98], Kirk and Stanton performed a study on three different gestures' formats. The research compared the performance of unmediated hands only, hands and sketch, and digital sketch only.

We divide the following section into two subsections. In the first, we present studies on pointers and annotations, and in the second, we look at behavior detection.

5.1. Pointers and Annotations

One of the first methods applied in remote assistance was a reality-augmenting telepointer controlled remotely to guide the local user [26]. The results of this approach showed that a remote user could effectively guide and direct a local user's activities. The participants of the study conducted in [99] for remote construction of a robot valued the pointing device because it helped them to make references to objects faster than they could with video.

The visible light path laser projector (VLLP) described in [100] was equipped with a laser projector and a mist generator. This solution had the advantage that the laser projector, VLLP, could instruct with not only a laser spot but also with simple line drawings.

A new setup was presented in [101], where a remote expert could annotate, point, and draw on the local worker's objects using a laser projector. At the time that it was developed, this method was intended to support pre- and post-surgical consultations between a surgeon and a remotely located patient.

A prototype system that enables a mobile worker to receive guidance from a remote helper in real time using freehand sketches projected directly onto the worker's environment is presented in [31].

A study comparison of the use of pointers and annotations in live video and still images was performed in [102]. The results showed that users collaborate more efficiently

using annotation cues than pointer cues for communicating object position and orientation information.

H. Jo [103] showed how Chili, a mobile phone call system, can provide on-video drawing capabilities and control of the viewpoint.

S. Kim proved, in his research on augmented visual communication cues [104], that both pointers and annotation could improve the feeling of being more connected, understanding the remote partner, and being together.

StickyLight [105] is a system based on a pico projector that allows a worker to draw annotations in the real local user's environment. The expert can draw and point via his tablet, which is connected to the local user's PC. The pico projector is paired with the local user's computer and performs the actions that the remote user is drawing.

Fakourfar et al., in their study about stabilized annotations [106], found that, by freezing the video, the temporal stabilization of the annotations was efficient if the camera did not have a fixed position. In this scenario, it could be difficult for the remote collaborator to return to live video if the viewed perspective was changed. Stabilized annotations proved to be useful when referencing objects but not when pointing or simulating object manipulation. The research found that, overall, users preferred stabilized annotations, even though they did not outperform non-stabilized annotations in all tasks.

In [107], a user was able to navigate in a virtual space in order to collaborate with other users, making use of 2D gestures in 3D reconstructed scenes. This method can be applied in virtual and augmented reality.

Rice et al., in their study [108], proposed a remote assistance platform that gave the remote user the ability to use video and virtual annotations to give explanations to the other user. The platform was compared with an instant messaging application that could be used to send text instructions for each captured image. The results showed that visual annotations are valuable when objects have a similar appearance.

AlphaRead [109] is a tool developed to reference objects by annotating them in a remote collaboration context. The system introduces the feature of object tracking, a solution for cases when the objects or the camera are moving. Each object has a readable label that the users can read in order to give instructions. This study showed that users found the object annotations useful and readable.

5.2. Behavior Detection

In this section, we look at studies that focused on the behavior of the human body in a remote collaboration context.

Fussel [27] proposed a method where the helper could follow a local worker's gaze through an eye tracker and head-mounted camera worn by the worker. However, a problem with this system was identified. Users could not make eye contact, and gaze awareness was lost. To improve this aspect, a conventional telepresence system was developed in [110], based on a see-through display that gave local and remote users the sensation that they were separated only by a vertical sheet of glass. This offered the ability to make eye contact and use non-verbal communication such as gaze and gestures.

As augmented reality is starting to be used more often, researchers are attempting to use it to overlay hands on a video stream. In [111], a solution is presented allowing hand gestures to be identified using augmented reality, and this method enables intuitive interaction during real-time computer vision.

HandsOnVideo [112] gives the helpers the possibility to point to objects using their hands. The remote worker's hands are overlaid onto the video stream on which the local user is able to see the instructions. An improved solution of HandsOnVideo is proposed in [63], which introduces hand gestures in a 3D context and allows the remote helper to use his hands directly to gesture. A special camera is used to capture hand movements in 3D.

BeThere [113] is based on a mobile smartphone configuration and depth sensors that allow two remote users to perform 3D mobile collaboration. This solution allows the remote user to perform 3D gestures in a shared environment and navigate in the 3D shared environment. In order to direct the user's attention, the system uses awareness cues such as helping hands and 3D annotations.

In his research, Zenati-Henda [114] showed how a gesture-recognizing module could detect and recognize the gestures of a remote user. The system recorded the position and the description of the gesture and displayed them as virtual hands on the local user's display.

In their study, Gupta et al. [115] tried to evaluate whether sharing the eye-tracking information of the local worker could improve performance during a remote assistance task. The study combined an eye tracker and a pointer and showed that the cues could enhance the feeling of co-presence between remote users.

Another approach to using both the eyes and the hands was proposed by Higuch et al. [116] in their research. The study addressed the problem of eye fixation during a remote guidance operation. Helpers used their eyes to identify the object of interest, and the hands were projected onto the local user's space to provide instructions for the manipulation of the objects.

A study by Li et al. [117] explored how shared gaze awareness impacted remote collaboration. The results showed that shared gaze information could be disruptive but improved remote coordination.

SharedSphere [90,91] is a prototype in which both the local and the remote user's hand gestures are captured and displayed on head-mounted devices worn by the users. The local user's hand gestures are captured by a video camera, while those of the remote user are captured by a hand-tracking sensor.

The study in [118] researched how gaze sharing in both directions influenced collaboration and communication between two remote participants. An eye tracker module was placed on the HMD of the local user and on the display of the remote helper. Results for most of the users showed that the system improved their awareness of their partner's focus, while some users mentioned that they could determine what the next step for their job execution would be.

HandsInTouch [119] combines gesture sharing with sketches in order to collaborate remotely. The study shows that hand gestures might be sufficient when performing easy tasks, but sketches together with gestures give better results when used for complex physical tasks involving objects.

Otsuki et al. [120], in their study, focused on how gaze cues support remote collaboration when using the ThirdEye display. ThirdEye is a hemispherical display that helps the local user to determine the gaze direction of the remote helper, which is then displayed on a mobile terminal. The experiment showed that the local user's attention is driven to the objects of interest when using ThirdEye faster than when not using it.

Omnigaze [121], another telepresence system, uses an omnidirectional video camera that has a spherical display on top of it. An eye tracker positioned at the remote user's location captures their gaze direction, which is then represented as information on the spherical display. This system enhanced the remote collaboration to some extent; however, it had many disadvantages, one being that the remote user could see the local user but not vice versa, which caused discomfort.

Wearable RemoteFusion [122] shares eye gaze and hand gestures in an MR environment. The local user can see in AR the hand gestures of the remote user guiding him, while the remote user has a view of the local worker's eye gaze in VR. The study showed that users had an above-average feeling of co-presence, while the remote user was under less psychical load and concentration.

Wang et al. [123] investigated whether sharing the eye gaze or the head pointer improved performance when resolving a task between two remote collaborators. The results showed that there were no significant differences between the two methods. Thus,

the head pointer method, which is less expensive, could be used instead of the eye tracking one.

ZoomTouch [124] is a system that allows multiple users to control a robot using a hand-tracking module remotely. Recognized hand gestures are used to control the robot, which has embedded tactile sensors.

Xiao et al. [125] reviewed eye-tracking prototypes categorized by their functionalities, subject used, and physical task types.

5.3. Combining Visual Communication Cues

The use of annotations and body cues has been studied separately, but there are several other studies that focus on the mixed use of visual communication cues, both related to body gestures and annotations.

Teo et al. [126] studied the use of hand gestures and annotations overlaid on a live 360-degree video. The system was based on MR; the local user wore an AR HMD, and the remote user wore a VR HMD. The remote user's gestures were tracked using a sensor and displayed in a 360-degree video feed as virtual hands. Besides the hand gesture cues, the user was allowed to use annotations, and his pointing gestures were captured and displayed as a ray pointer as seen through the local user's AR device. The annotations had fixed spots both in the virtual and the real environment, even if the local user's head direction changed. The study showed that the participants could finish the task faster and understand the remote instructions better when using visual annotations.

Kim et al. [127] researched the effect of combining visual communication cues such as hand gestures, sketches, and pointing in a mixed reality remote collaboration. The study showed that using sketches and hand gestures gave the best results regarding task performance, but the pointing cues did not improve the performance significantly. Moreover, the sketches and the pointing cues involved higher mental effort for the users.

The research of Teo et al. [128,129] focused on proposing communication cues when using an MR system based on a 360-degree panorama reconstructed in a 3D environment. The remote user could control a virtual ray pointer using hand gestures or add sketches (drawings), which were displayed through the AR HMD of the local user and overlaid onto the 360-degree panoramas seen by the remote user.

Bai et al. [130] modeled the eye gaze of the remote user as a virtual ray cast line and the remote user's hand as a 3D mesh; both were overlaid onto the local user's AR view. In order to help the users to identify their partner's location and viewing direction, the authors proposed the use of a 3D arrow cue and a virtual avatar as a virtual head frustum that points to the location of the other user. The head frustum represents the other user's head direction, while the 3D arrow is modeled as a pin arrow pointing to the other user's avatar head. The results showed that when combining hand gestures and eye gaze, the task completion time was reduced, and the users had a better feeling of co-presence. For the local user, the required mental effort was considered lower when using the gesture cues, but both users said that they preferred combined cues over gesture or gaze alone.

In Table 3, we classify the publications that address the use of communication cues in remote collaboration and assistance by the application and the communication method employed to accomplish the given task. In this case, the synthesis in Table 3 allows the straightforward identification of scientific papers that present a method applied in a certain field. With the advent of new 5G communications technologies, it is possible that researchers' efforts will focus on such methods. Moreover, many more applications will be developed, and the global state of the pandemic will accelerate research in these areas.

Surveying the scientific literature, it can be observed that efficient communication presupposes sound, image, and orientation towards the object being discussed. Digitization together with remote assistance and collaboration to accomplish tasks will be the future challenges for technological development.

Table 3. Communication cues.

Application (Task Type)	Communication Cues			
	Annotations	**Gestures**	**Pointing**	**Eye Gaze**
Robot Building		[96]	[99]	
Object Identification	[103,107,128]	[128]	[100]	[121]
Surgery	[101]			
Lego Assembly	[97,98,102,105,106, 109,122,127]	[97,98,113,116,119, 122,123,127]	[102,115]	[115,116,122]
Puzzle Assembly	[104]		[104]	
Wiring Assembly	[108]			
Education		[111]		[110]
Industry		[112]		
Printer Assembly		[114]		
Design				[124]
Objects Manipulation	[126,129]	[126,129,130]		[120]
Physical Task		[125]		

6. Discussions

In this literature review, we identify the main applications where remote assistance is used. Table 4 represents the type of engineering applications described and utilized as case studies throughout this survey.

Table 4. Applications of remote collaboration system.

Applications	Article Reference
Healthcare	[1–10,61,101]
Industry	[13–21,30,31,36,38,40,41,92,108,112]
Education	[11,12,26,29,34,39,47,53,95,96,99,110,111,121,125]
Assembly Task	[28,35,37,42,48–52,57,62,63,66–73,97,98,102,104– 106,109,113–117,119,122,123,127]
Aviation	[32]
Crime Investigation	[22]
Maintenance	[27,60,65,74–78]
Object Identification	[43–46,55,56,58,64,71,88–91,100,103,107,128]
Objects Manipulation	[54,79–87,93,94,121,126,129,130]
Design	[59,118,124]

To summarize the current literature review, as seen in Figure 6, we depict the trend for devices used for displaying the working context, and the development of the industry is surveyed during the periods 1999–2006, 2007–2010, 2011–2015, and 2016–2020. We observe an increase in the production of device types over all periods for all the studied technologies—HMCs, HMDs, video cameras, and smart devices and displays—with the exception of robotic devices and projectors, which are not as widely examined over time, perhaps because of their complexity and automation needs. We also observe a decrease in the use of discrete video cameras and displays that are integrated into HMCs and HMDs. The tendencies are towards the smallest form factor of the product; HMDs were the most dominant solution for the period 2016–2020. The interest in LCD displays and video cameras decreased from 2011 to 2015 and from 2016 to 2020. There were few identified HMDs and HMCs during 1999–2006, and the reason is that the integration and the miniaturization techniques were very incipient. We can state that the emergence of these devices, HMCs, HMDs, smart devices (tablet/phone/remote control), robotic devices, and

projectors, started with a proof of concept, and the demonstration of their feasibility persuaded the manufacturing companies that the chosen solution was reliable.
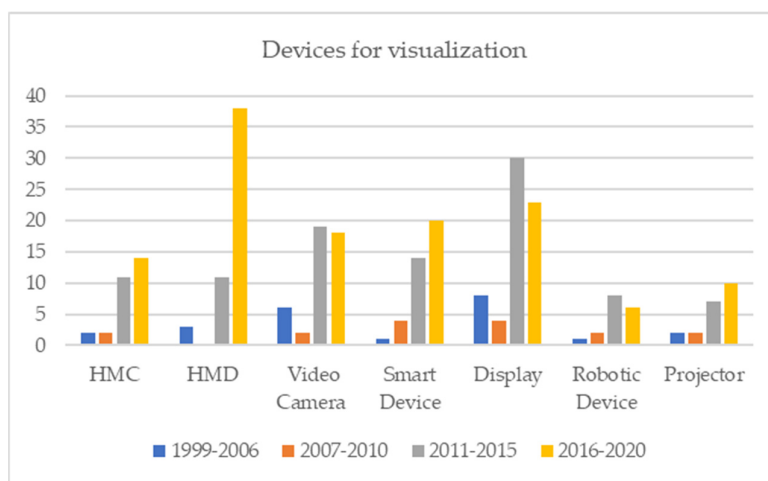

Figure 6. Number and type of devices for capturing and visualization between 1999 and 2020.

The study of the occurrence of communication cues in the same periods of time is presented in Figure 7. We can see that 20 years ago, pointing was used more often as a communication cue when collaborating remotely; in the last four years, the emphasis has been on annotations, gestures, and eye gaze.
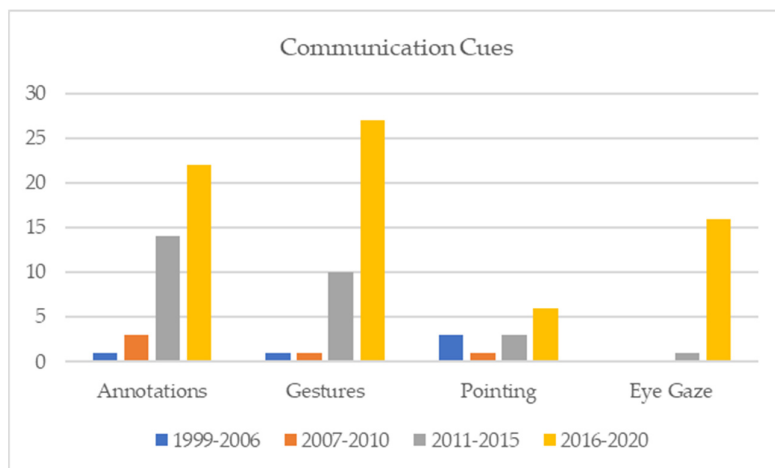

Figure 7. Communication cues occurrences for the period 1999–2020.

Overall, studies showed that, during a remote collaboration, tasks are completed better and faster when the visual workspace is shared between the participants.

7. Conclusions

From the remote collaboration systems presented above, several issues linked to the devices and video cameras can be specified:

- Using fixed cameras can reduce the field of view for the remote specialist and make it difficult to determine where the local user is looking.
- Head-mounted cameras (HMCs), besides being uncomfortable for the wearer, restrict the field of view for the helper, who can only see what the worker is seeing. On the other hand, HMCs allow the worker to use their hands to perform tasks.

- Handheld devices limit a local user's ability to use both hands freely, but the research shows that they are a good option when working on static tasks.
- Head-mounted devices proved to be suitable for dynamic tasks but need stabilization techniques for captured video because the images are shaky, and viewers become dizzy.
- An independent view for the remote participant improves the viewer's confidence significantly, and verbal communication is reduced used during the collaboration.
- Smartphone applications used in remote collaboration still have limited performance because of the heavy use of networks when streaming video.
- No published works were found on the extent to which more experts from different geographically distant locations can offer collaborative assistance for a machine/process, etc.

When comparing the visual display methods, mixed reality proved to be very useful for tasks that needed complex design instructions and decreased time and mental effort. Reconstructing 3D scenes is challenging because they often have to be dynamically updated, requiring higher processing power.

Regarding communication cues, pointing was the best method to use when a user wished to be quick and precise when indicating an object. Annotation cues such as sketches proved to be useful, providing spatial information when manipulating objects. Hand gestures could express more information, such as pointing, emotion, appreciation, and shapes. Researchers found that a user finished faster a task when using hand gestures than without them.

This study reviewed different prototypes, systems, and methods used in remote collaboration over the last two decades. We categorized and analyzed the systems and functions of the devices used in remote collaboration, the display view of the local environment, and the communication cues used between remote collaborators. We discussed the limitations and disadvantages of the current system components and methods to better understand why these kinds of systems are not used at a large scale in industry and to help to make future decisions when designing a remote collaboration system. Table 5 summarizes the benefits and the drawbacks of the devices used in remote collaboration.

Table 5. Applications of remote collaboration systems.

Component	Benefits	Drawbacks
Fixed Camera	Low cost	Reduced FOV [1]
Head-Mounted Camera	Worker's hands are free for tasks	Reduced FOV Comfortless
Head-Mounted Device	Suitable for dynamic tasks Close to reality—immersive view	High Costs Comfortless
Smart Device [2]	Suitable for static tasks Low cost Plug and play Connect anywhere, anytime	Poor network performance Hands occupied Reduced costs
Robotic Device	Worker's hands are free for tasks High degree of automation and control	Complexity Engineering costs Materials costs

[1] FOV = field of view; [2] Smart Device = smartphone/tablet.

Remote assistance began to be used in surgery more than 20 years ago, and it is now beginning to be employed in other fields, such as education or industry.

We have seen that the latest technologies used in remote collaboration are those that allow the remote user to have an independent view of the local user's environment and

offer the capability of a 360-degree panorama view. These systems employ HMDs together with MR, AR, or VR. Smart devices such as mobile phones and tablets, even if they are more accessible than HMDs and have better capabilities (such as video streaming), are not employed at a large scale for remote assistance.

It can be concluded that there are many challenges in obtaining a remote collaboration system that allows the remote users to have a controlled view and make them feel as if they are immersed within the local user's environment, together with the capability of transmitting instructions in a natural manner. Digitization, together with remote assistance and collaboration to accomplish tasks, will be the future challenges for technological development. Due to the new 5G communications technologies, researchers' efforts will certainly focus on such methods. Moreover, many more applications will be developed, and the problems caused by the pandemic situation will accelerate research in these areas.

Author Contributions: Conceptualization, R.D. and I.S.; methodology, R.D., C.D., and I.S.; validation, R.D., C.D., P.N., and I.S.; formal analysis, R.D.; investigation, R.D.; writing—original draft preparation, R.D. and C.D.; writing—review and editing, R.D, C.D., P.N., and I.S.; supervision, I.S.; project administration, I.S.; funding acquisition, I.S. All authors have read and agreed to the published version of the manuscript.

References

1. Villarruel, J.E.G.; Corona, B.T. Proposal for a Remote Surgery System Based on Wireless Communications, Electromyography and Robotics. In Proceedings of the 2008 Electronics, Robotics and Automotive Mechanics Conference (CERMA '08), Cuernavaca, Mexico, 30 September–3 October 2008; pp. 93–98.
2. Zemmar, A.; Lozano, A.M.; Nelson, B.J. The rise of robots in surgical environments during COVID-19. *Nat. Mach. Intell.* **2020**, *2*, 566–572, doi:10.1038/s42256-020-00238-2.
3. Zineddine, M.; Arafa, N. Attitude towards Robot Assisted Surgery: UAE context. In Proceedings of the 2013 9th International Conference on Innovations in Information Technology (IIT), Al Ain, United Arab Emirates, 17—19 March 2013; pp. 175–179.
4. Suthakorn, J. A concept on Cooperative Tele-Surgical System based on Image-Guiding and robotic technology. In Proceedings of the 2012 Pan American Health Care Exchanges, Miami, FL, USA, 26—31 March 2012; pp. 41–45, doi:10.1109/pahce.2012.6233437.
5. Ma, X.; Guo, S.; Xiao, N.; Gao, B.; Guo, J.; Tamiya, T.; Kawanishi, M.; Kawanishi, M. Remote catheterization using a new robotic catheter manipulating system. In Proceedings of the 2013 ICME International Conference on Complex Medical Engineering, Beijing, China, 25–28 May 2013; pp. 394–398.
6. Yanhong, F.; B in, W.; Fengjuan, H.; Wenqiang, T. Research on teleoperation surgery simulation system based on virtual reality. In Proceeding of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 5830–5834.
7. Cardenas, I.S.; Kim, J.-H.; I.S., C. Advanced Technique for Tele-operated Surgery Using an Intelligent Head-mount Display System. In Proceedings of the 2013 29th Southern Biomedical Engineering Conference, Miami, FL, USA, 3–5 May 2013; pp. 143–144.
8. Dardona, T.; Eslamian, S.; Reisner, L.A.; Pandya, A. Remote Presence: Development and Usability Evaluation of a Head-Mounted Display for Camera Control on the da Vinci Surgical System. *Robotics* **2019**, *8*, 31, doi:10.3390/robotics8020031.
9. Colan, J.; Nakanishi, J.; Aoyama, T.; Hasegawa, Y. A Cooperative Human-Robot Interface for Constrained Manipulation in Robot-Assisted Endonasal Surgery. *Appl. Sci.* **2020**, *10*, 4809, doi:10.3390/app10144809.
10. Widick, L.; Talkington, J.; Bajwa, G.; Dantu, R.; Logan, W. A framework for secured collaboration in mHealth. In Proceedings of the 2015 International Conference on Collaboration Technologies and Systems (CTS), Atlanta, GA, USA, 1–5 June 2015; pp. 364–369.
11. Cidota, M.A.; Datcu, D. Remote Assistance System in Augmented Reality for Early School Dropout Prevention. In Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Recife, Brazil, 1–9 Nov. 2020; pp. 321–325.

12. Gu, J.; Li, N.; Duh, H.B.-L. A remote mobile collaborative AR system for learning in physics. In Proceedings of the 2011 IEEE Virtual Reality Conference, Singapore, 19–23 March 2011; pp. 257–258.
13. Bednarz, T.; James, C.; Caris, C.; Haustein, K.; Adcock, M.; Gunn, C. Applications of networked virtual reality for tele-operation and tele-assistance systems in the mining industry. In Proceedings of the 10th International Conference on Predictive Models in Software Engineering, Hong Kong, China, 11 December 2011; pp. 459–462.
14. Back, M.; Childs, T.; Dunnigan, A.; Foote, J.; Gattepally, S.; Liew, B.; Shingu, J.; Vaughan, J. The virtual factory: Exploring 3D worlds as industrial collaboration and control environments. In Proceedings of the 2010 IEEE Virtual Reality Conference (VR), Boston, MA, USA, 20–24 March 2010; pp. 257–258.
15. Back, M.; Kimber, D.; Rieffel, E.; Dunnigan, A.; Liew, B.; Gattepally, S.; Foote, J.; Shingu, J.; Vaughan, J. The virtual chocolate factory: Building a real world mixed-reality system for industrial collaboration and control. In Proceedings of the 2010 IEEE International Conference on Multimedia and Expo, Singapore, 19–23 July 2010; pp. 1160–1165.
16. Kesavan, S. State-of-Art, Modular, Cloud Enabled Video Collaboration Solution for Next Generation Manufacturing Industries. In Proceedings of the 2019 Global Conference for Advancement in Technology (GCAT), Bangalore, India, 18–20 October 2019; pp. 1–6.
17. Blog, W.E. Introducing Two New Mixed Reality Business Applications: Microsoft Remote Assist and Microsoft Layout. Windows Experience Blog, 7 May 2018. Available online: https://blogs.windows.com/windowsexperience/2018/05/07/introducing-two-new-mixed-reality-business-applications-microsoft-remote-assist-and-microsoft-layout/ (accessed on 28 April 2021).
18. HoloLens 2—Overview, Features, and Specs|Microsoft HoloLens. Available online: https://www.microsoft.com/ro-ro/hololens/hardware (accessed on 2 July 2021).
19. Dynamics 365 Remote Assist Mobile Overview—Dynamics 365 Mixed Reality. Available online: https://docs.microsoft.com/en-us/dynamics365/mixed-reality/remote-assist/mobile-app/remote-assist-mobile-overview (accessed on 28 April 2021).
20. Campero-Jurado, I.; Márquez-Sánchez, S.; Quintanar-Gómez, J.; Rodríguez, S.; Corchado, J. Smart Helmet 5.0 for Industrial Internet of Things Using Artificial Intelligence. *Sensors* **2020**, *20*, 6241, doi:10.3390/s20216241.
21. Hitachi and Microsoft form a Strategic Alliance to Advance Next-Generation Digital Solutions for Manufacturing and Logistics Across Southeast Asia, North America and Japan. Available online: https://www.hitachi.com/New/cnews/month/2020/06/200626.pdf (accessed on 28 April 2021).
22. Poelman, R.; Akman, O.; Lukosch, S.; Jonker, P. As if Being There: Mediated Reality for Crime Scene Investigation. In Proceedings of the CSCW '12: 2012 ACM Conference on Computer Supported Cooperative Work, New York, NY, USA, 11 February 2012, pp.1267–1276.
23. IEEE Xplore. Available online: https://ieeexplore.ieee.org/ (accessed on 12 September 2021).
24. Yang, K.; Setubal, J.C. A Whole Genome Simulator of Prokaryote Genome Evolution. In Proceedings of the Second ACM Conference on Bioinformatics, Computational Biology and Biomedicine, Chicago, IL, USA, 31 July–3 August 2011; Grossman, R., Rzhetsky, A., Eds.; ACM: New York, NY, USA; pp. 508–510. Available online: http://dl.acm.org (accessed on 10 July 2012).
25. Google Trends. Available online: https://trends.google.com/trends/explore?date=2004-01-01%202021-08-01&q=remote%20collaboration,remote%20team%20collaboration,remote%20intervention (accessed on 12 September 2021).
26. Bauer, M.; Kortuem, G.; Segall, Z. "Where are you pointing at?" A study of remote collaboration in a wearable videoconference system. In *Digest of Papers, Proceedings of the Third International Symposium on Wearable Computers*; Institute of Electrical and Electronics Engineers (IEEE), San Francisco, CA, USA,18–19 Oct. 2003; pp. 151–158.
27. Fussell, S.R.; Setlock, L.D.; Kraut, R.E. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In Proceedings of the Conference on Human Factors in Computing Systems—CHI '03, Ft. Lauderdale, Florida, USA, 5 April 2003; pp. 513–520.
28. Kurata, T.; Sakata, N.; Kourogi, M.; Kuzuoka, H.; Billinghurst, M. Remote Collaboration using a Shoulder-Worn Active Camera/Laser. *Eighth Int. Symp. Wearable Comput.* **2004**, *1*, 62–69, doi:10.1109/iswc.2004.37.
29. Machino, T.; Nanjo, Y.; Yanagihara, Y.; Kawata, H.; Iwaki, S.; Shimokura, K. Robot-augmented communication: a remote-collaboration system based on a shared field of view in real space. In Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, 2–6 August 2005; pp. 2203–2209.
30. Machino, T.; Iwaki, S.; Kawata, H.; Yanagihara, Y.; Nanjo, Y.; Shimokura, K.-I. Remote-collaboration system using mobile robot with camera and projector. In Proceedings of the 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006, Orlando, FL, USA, 15–19 May 2006; pp. 4063–4068.
31. Adcock, M.; Gunn, C. Annotating with 'sticky' light for remote guidance. In *ACM SIGGRAPH ASIA 2010 Posters on—SA '10*; Association for Computing Machinery, Seoul, Republic of Korea, 15 December 2010; p. 1.
32. Gauglitz, S.; Lee, C.; Turk, M.; Höllerer, T. Integrating the physical environment into mobile remote collaboration. In Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '12, San Francisco, California, USA, 21 September 2012; pp. 241–250.
33. Gauglitz, S.; Nuernberger, B.; Turk, M.; Höllerer, T. World-stabilized annotations and virtual scene navigation for remote collaboration. In Proceedings of the 27th annual ACM Symposium on User Interface Software and Technology, Honolulu, Hawaii, USA, 5 October 2014; pp. 449–459.
34. Gurevich, P.; Lanir, J.; Cohen, B.; Stone, R. TeleAdvisor: A versatile augmented reality tool for remote assistance. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA, 5 May 2012; pp. 619–622.

35. Huang, W.; Alem, L. HandsinAir: A Wearable System for Remote Collaboration on Physical Tasks. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion–CSCW '13, Association for Computing Machinery, New York, NY, USA, 23 February 2013; pp. 153–156.

36. Robert, K.; Zhu, D.; Huang, W.; Alem, L.; Gedeon, T. MobileHelper: Remote Guiding Using Smart Mobile Devices, Hand Gestures and Augmented Reality. In Proceedings of the SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications on—SA '13, Association for Computing Machinery, New York, NY, USA, 19 November 2013; Volume 39; p. 39.

37. Tait, M.; Tsai, T.; Sakata, N.; Billinghurst, M.; Vartiainen, E. A projected augmented reality system for remote collaboration. In Proceedings of the 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Adelaide, Australia, 1–4 Oct. 2013; pp. 1–6, doi:10.1109/ismar.2013.6671838.

38. Domova, V.; Vartiainen, E.; Englund, M. Designing a Remote Video Collaboration System for Industrial Settings. In Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces–ITS '14, Association for Computing Machinery, New York, 16 November 2014; pp. 229–238.

39. Onishi, Y.; Tanaka, K.; Nakanishi, H.; Y., O.; K., T.; H., N. PopArm: A robot arm for embodying video-mediated pointing behaviors. In Proceedings of the 2014 International Conference on Collaboration Technologies and Systems (CTS), Minneapolis, MN, USA, 19–23 May 2014; pp. 137–141.

40. Benavides, X.; Amores, J.; Maes, P. Remot-IO: A System for Reaching into the Environment of a Remote Collaborator, In Adjunct. Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, Association for Computing Machinery, New York, NY, USA, 06 November 2015; pp. 99–100, doi:10.1145/2815585.2815738.

41. Vartiainen, E.; Domova, V.; Englund, M. Expert on Wheels: An Approach to Remote Collaboration. In Proceedings of the 3rd International Conference on Human-Agent Interaction, Association for Computing Machinery, New York, NY, USA 21 October 2015; pp. 49–54.

42. Johnson, S.; Gibson, M.; Mutlu, B. Handheld or Handsfree? Remote Collaboration via Lightweight Head-Mounted Dis-plays and Handheld Devices. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, Vancouver, BC, Canada, 28 February 2015; pp. 1825–1836.

43. Johnson, S.; Rae, I.; Mutlu, B.; Takayama, L. Can You See Me Now? How Field of View Affects Collaboration in Robotic Telepresence. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, 18 April 2015; pp. 2397–2406.

44. Kasahara, S.; Nagai, S.; Rekimoto, J. JackIn Head: Immersive Visual Telepresence System with Omnidirectional Wearable Camera. *IEEE Trans. Vis. Comput. Graph.* **2016**, *23*, 1222–1234, doi:10.1109/TVCG.2016.2642947.

45. Tait, M.; Billinghurst, M. The Effect of View Independence in a Collaborative AR System. *Comput. Support. Cooperative Work. (CSCW)* **2015**, *24*, 563–589, doi:10.1007/s10606-015-9231-8.

46. Ahn, J.-G.; Kim, G.J. Remote collaboration using a tele-presence mobile projector robot tele-operated by a smartphone. In Proceedings of the 2016 IEEE/SICE International Symposium on System Integration (SII), Sapporo, Japan, 13-15 Dec. 2016; pp. 236–241, doi:10.1109/sii.2016.7844004.

47. Essig, K.; Strenge, B.; Schack, T. ADAMAAS: Towards Smart Glasses for Mobile and Personalized Action Assistance. In Proceedings of the 9th ACM International Conference on Distributed Event-Based Systems, Association for Computing Machinery, New York, NY, USA, 29 June 2016; p. 46.

48. Kratz, S.; Ferriera, F.R. Immersed remotely: Evaluating the use of Head Mounted Devices for remote collaboration in robotic telepresence. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), New York, NY, USA, 26–31 August 2016; pp. 638–645.

49. Lee, Y.; Masai, K.; Kunze, K.; Sugimoto, M.; Billinghurst, M. A Remote Collaboration System with Empathy Glasses. In Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Merida, Mexico 19–23 September 2016; pp. 342–343.

50. Masai, K.; Kunze, K.; Sugimoto, M.; Billinghurst, M. Empathy Glasses. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, San Jose, California, USA, 7 May 2016; pp. 1257–1263.

51. Komiyama, R.; Miyaki, T.; Rekimoto, J. JackIn space: Designing a seamless transition between first and third person view for effective telepresence collaborations. In Proceedings of the 8th Augmented Human International Conference, Silicon Valley, California, USA, 16 March 2017; p. 14.

52. Le, K.-D.; Zhu, K.; Fjeld, M. Mirrortablet: Exploring a low-cost mobile system for capturing unmediated hand gestures in remote collaboration. In Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, Association for Computing Machinery, New York, NY, USA, 26 November 2017; pp. 79–89.

53. Tobita, H. Gutsy-Avatar: Computational Assimilation for Advanced Communication and Collaboration. In Proceedings of the 2017 First IEEE International Conference on Robotic Computing (IRC), Taichung, Taiwan, 10–12 April 2017; pp. 8–13.

54. Feick, M.; Mok, T.; Tang, A.; Oehlberg, L.; Sharlin, E. Perspective on and Re-orientation of Physical Proxies in Object-Focused Remote Collaboration. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal QC, Canada, Association for Computing Machinery, New York, NY, USA 21 April 2018; p. 281.

55. Lee, G.A.; Teo, T.; Kim, S.; Billinghurst, M. A User Study on MR Remote Collaboration Using Live 360 Video. In Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 16–20 October 2018; pp. 153–164.

56. Unver, B.; D'Angelo, S.; Miller, M.; Tang, J.C.; Venolia, G.; Inkpen, K. Hands-Free Remote Collaboration Over Video: Exploring Viewer and Streamer Reactions. In Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces, Tokyo, Japan, 19 November 2018; pp. 85–95.

57. Kwon, J.U.; Hwang, J.-I.; Park, J.; Ahn, S.C. Fully Asymmetric Remote Collaboration System. *IEEE Access* **2019**, *7*, 54155–54166, doi:10.1109/access.2019.2912397.

58. Teo, T.; Lee, G.; Billinghurst, M.; Adcock, M. Merging Live and Static 360 Panoramas Inside a 3D Scene for Mixed Reality Remote Collaboration. In Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Beijing, China, 10–18 October 2019; pp. 22–25.

59. Li, Z.; Teo, T.; Chan, L.; Lee, G.; Adcock, M.; Billinghurst, M.; Koike, H. OmniGlobeVR. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, Association for Computing Machinery, New York, NY, USA, 3 July 2020; pp. 615–625.

60. Mohr, P.; Mori, S.; Langlotz, T.; Thomas, B.H.; Schmalstieg, D.; Kalkofen, D. Mixed Reality Light Fields for Interactive Remote Assistance. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, 21 April 2020; pp. 1–12.

61. Welch, G.; Sonnenwald, D.H.; Mayer-Patel, K.; Yang, R.; State, A.; Towles, H.; Cairns, B.; Fuchs, H. Remote 3D Medical Consultation. In *Virtual Realities: Dagstuhl Seminar 2008*; Brunnett, G., Coquillart, S., Welch, G., Eds.; Springer: Vienna, Austria, 2011; pp. 139–159.

62. Kainz, B.; Schmalstieg, D.; Hauswiesner, S.; Reitmayr, G.; Steinberger, M.; Grasset, R.; Gruber, L.; Veas, E.; Kalkofen, D.; Seichter, H. OmniKinect. In Proceedings of the 18th ACM symposium on Virtual reality software and technology—VRST '12, Association for Computing Machinery, New York, NY, USA, 10 December 2012; pp. 25–32.

63. Tecchia, F.; Alem, L.; Huang, W. 3D helping hands: A gesture based MR system for remote collaboration. In Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry—VRCAI '12, Association for Computing Machinery, New York, NY, USA, 2 December 2012; pp. 323–328.

64. Adcock, M.; Anderson, S.; Thomas, B. RemoteFusion: Real time depth camera fusion for remote collaboration on physical tasks. In Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry—VRCAI '13, Hong Kong, Hong Kong, 17 November 2013; pp. 235–242.

65. Gauglitz, S.; Nuernberger, B.; Turk, M.; Höllerer, T. 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—VRST '14, Edinburgh, Scotland, 11 November 2014; pp. 197–205.

66. Gao, L.; Bai, H.; He, W.; Billinghurst, M.; Lindeman, R.W. Real-time visual representations for mobile mixed reality remote collaboration. In Proceedings of the SIGGRAPH Asia 2018 Virtual & Augmented Reality on—SA '18, Tokyo, Japan, 4 December 2018; pp. 1–2.

67. Oculus Rift. Available online: https://en.wikipedia.org/wiki/Oculus_Rift (accessed on 24 September 2021).

68. HTC Vive. Available online: https://en.wikipedia.org/wiki/HTC_Vive (accessed on 24 September 2021).

69. Oculus Quest. Available online: https://en.wikipedia.org/wiki/Oculus_Quest (accessed on 24 September 2021).

70. "Virtual Reality," Wikipedia. Available online: https://en.wikipedia.org/wiki/Virtual_reality (accessed on 24 September 2021).

71. Piumsomboon, T.; Day, A.; Ens, B.; Lee, Y.; Lee, G.; Billinghurst, M. Exploring enhancements for remote mixed reality collaboration. In Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on—SA '17, Bangkok, Thailand, November 27, 2017; p. 16.

72. Gao, L.; Bai, H.; Lee, G.; Billinghurst, M. An oriented point-cloud view for MR remote collaboration. In Proceedings of the SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications, Macau, 28 November 2016; pp. 8–8:4.

73. Gao, L.; Bai, H.; Lindeman, R.; Billinghurst, M. Static local environment capturing and sharing for MR remote collaboration. In Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on—SA '17, Bangkok, Thailand, 27 November 2017; p. 17.

74. Elvezio, C.; Sukan, M.; Oda, O.; Feiner, S.; Tversky, B. Remote collaboration in AR and VR using virtual replicas. In Proceedings of the ACM SIGGRAPH 2017 VR Village, Los Angeles, California, 30 July 2017; p. 13.

75. Van Krevelen, R. Augmented Reality: Technologies, Applications, and Limitations. 2007. https://doi.org/10.13140/RG.2.1.1874.7929 (accessed on 24 September 2021)

76. Bottecchia, S.; Cieutat, J.-M.; Jessel, J.-P. T.A.C. T.A.C: Augmented reality system for collaborative tele-assistance in the field of maintenance through internet. In Proceedings of the 1st Augmented Human International Conference, Megève, France, 2 April 2010; pp. 1–7.

77. Schneider, M.; Rambach, J.; Stricker, D. Augmented reality based on edge computing using the example of remote live support. In Proceedings of the 2017 IEEE International Conference on Industrial Technology (ICIT), Toronto, ON, Canada, 22–25 March 2017; pp. 1277–1282.

78. Zillner, J.; Mendez, E.; Wagner, D. Augmented Reality Remote Collaboration with Dense Reconstruction. In Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 16–20 October 2018; pp. 38–39.

79. Rasmussen, T.A.; Huang, W. SceneCam: Using AR to improve Multi-Camera Remote Collaboration. In Proceedings of the SIGGRAPH Asia 2019 XR, Brisbane, QLD, Australia, 17 November 2019; pp. 36–37.

80. Rasmussen, T.A.; Huang, W. SceneCam: Improving Multi-camera Remote Collaboration using Augmented Reality. In Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Beijing, China, 10–18 October 2019; pp. 28–33.

81. Yang, P.; Kitahara, I.; Ohta, Y. [POSTER] Remote Mixed Reality System Supporting Interactions with Virtualized Objects. In Proceedings of the 2015 IEEE International Symposium on Mixed and Augmented Reality; Institute of Electrical and Electronics Engineers (IEEE), 2015; pp. 64–67.

82. Müller, J.; Rädle, R.; Reiterer, H. Remote Collaboration with Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, Colorado, USA, 2 May 2017; pp. 6481–6486.

83. Feick, M.; Tang, A.; Bateman, S. Mixed-Reality for Object-Focused Remote Collaboration. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings, Berlin, Germany, 11 October 2018; pp. 63–65.

84. Teo, T.; Lee, G.; Billinghurst, M.; Adcock, M. 360Drops: Mixed Reality Remote Collaboration using 360 Panoramas within the 3D Scene*. In *SIGGRAPH Asia 2019 Emerging Technologies*, Brisbane, QLD, Australia, Association for Computing Machinery, New York, NY, USA, 17 November 2019; pp. 1–2.

85. Teo, T.; Hayati, A.F.; Lee, G.A.; Billinghurst, M.; Adcock, M. A Technique for Mixed Reality Remote Collaboration using 360 Panoramas in 3D Reconstructed Scenes. In Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology, Association for Computing Machinery, New York, NY, USA, 12 November 2019; p. 23.

86. Teo, T.; Lawrence, L.; Lee, G.A.; Billinghurst, M.; Adcock, M. Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, 2 May 2019; pp. 1–14.

87. Gao, L.; Bai, H.; Billinghurst, M.; Lindeman, R.W. User Behaviour Analysis of Mixed Reality Remote Collaboration with a Hybrid View Interface. In Proceedings of the 32nd Australian Conference on Human-Computer Interaction, Association for Computing Machinery New York, NY, USA, 2 December 2020; pp. 629–638.

88. Kasahara, S.; Nagai, S.; Rekimoto, J. LiveSphere: Immersive experience sharing with 360 degrees head-mounted cameras. In Proceedings of the adjunct publication of the 27th annual ACM symposium on User interface software and technology—UIST'14 Adjunct, Honolulu, Hawaii, USA, 5 October 2014; pp. 61–62.

89. Kasahara, S.; Rekimoto, J. JackIn head: Immersive visual telepresence system with omnidirectional wearable camera for remote collaboration. In Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology, Beijing, China, 13 November 2015; pp. 217–225.

90. Lee, G.A.; Teo, T.; Kim, S.; Billinghurst, M. Mixed reality collaboration through sharing a live panorama. In Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications on—SA '17, Bangkok, Thailand, 27 November 2017; p. 14.

91. Lee, G.A.; Teo, T.; Kim, S.; Billinghurst, M. Sharedsphere: MR collaboration through shared live panorama. In Proceedings of the SIGGRAPH Asia 2017 Emerging Technologies, Bangkok, Thailand, 27 November 2017; p. 12.

92. Kangas, J.; Sand, A.; Jokela, T.; Piippo, P.; Eskolin, P.; Salmimaa, M.; Raisamo, R. Remote Expert for Assistance in a Physical Operational Task. In Proceedings of the Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA, 20 April 2018; p. LBW127.

93. Speicher, M.; Cao, J.; Yu, A.; Zhang, H.; Nebeling, M. 360Anywhere: Mobile Ad-hoc Collaboration in Any Environment using 360 Video and Augmented Reality. *Proc. ACM Human-Computer Interact.* **2018**, *2*, 1–20, doi:10.1145/3229091.

94. Piumsomboon, T.; Lee, G.A.; Irlitti, A.; Ens, B.; Thomas, B.H.; Billinghurst, M. On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, 2 May 2019; p. 228.

95. Fussell, S.R.; Kraut, R.E.; Siegel, J. Coordination of communication: effects of shared visual context on collaborative work. In Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, Philadelphia, PA, USA, 1 December 2000; pp. 21–30, doi:10.1145/358916.358947S.

96. Fussell, S.R.; Setlock, L.D.; Yang, J.; Ou, J.; Mauer, E.; Kramer, A.D.I. Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks. *Human–Comput. Interact.* **2004**, *19*, 273–309, doi:10.1207/s15327051hci1903_3.

97. Kirk, D.; Rodden, T.; Fraser, D.S. Turn it this way: Grounding collaborative action with remote gestures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA, 29 April 2007; pp. 1039–1048.

98. Kirk, D.; Fraser, D.S. Comparing remote gesture technologies for supporting collaborative physical tasks. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, Québec, Canada, 22 April 2006; pp. 1191–1200.

99. Fussell, S.R.; Setlock, L.D.; Parker, E.M.; Yang, J. Assessing the value of a cursor pointing device for remote collaboration on physical tasks. In Proceedings of the CHI '03 Extended Abstracts on Human Factors in Computing Systems—CHI '03, New York, NY, USA, 5 April 2003; pp. 788–789.

100. Sakata, N.; Okamoto, S.; Nishida, S. Applications of Visible Light Path Laser Projector. Presented at the Symposium on Human Interface, Springer, Berlin, Heidelberg, 19 July 2009, pp. 611–618. https://doi.org/10.1007/978-3-642-02556-3_69

101. Palmer, D.; Adcock, M.; Smith, J.; Hutchins, M.; Gunn, C.; Stevenson, D.; Taylor, K. Annotating with light for remote guidance. In Proceedings of the 2007 Conference of the Computer-Human Interaction Special Interest Group (CHISIG) of Australia on Computer-Human Interaction: Design: Activities, Artifacts and Environments–OZCHI '07, Adelaide, Australia, 28 November 2007; pp. 103–110.

102. Kim, S.; Lee, G.A.; Sakata, N.; Dunser, A.; Vartiainen, E.; Billinghurst, M. Study of augmented gesture communication cues and view sharing in remote collaboration. In Proceedings of the 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Adelaide, SA, Australia, 1–4 October 2013; pp. 261–262.

103. Jo, H.; Hwang, S. Chili: Viewpoint control and on-video drawing for mobile video calls. In Proceedings of the CHI '13 Extended Abstracts on Human Factors in Computing Systems on—CHI EA '13, Paris, France, 27 April 2013; pp. 1425–1430.

104. Kim, S.; Lee, G.; Sakata, N.; Billinghurst, M. Improving co-presence with augmented visual communication cues for sharing experience through video conference. In Proceedings of the 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 10–12 September 2014; pp. 83–92.

105. Adcock, M.; Gunn, C. Using Projected Light for Mobile Remote Guidance. *Comput. Support. Cooperative Work.* **2015**, *24*, 591–611, doi:10.1007/s10606-015-9237-2.

106. Fakourfar, O.; Ta, K.; Tang, R.; Bateman, S.; Tang, A. Stabilized Annotations for Mobile Remote Assistance. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, California, USA, 7 May 2016; pp. 1548–1560.

107. Nuernberger, B.; Lien, K.-C.; Grinta, L.; Sweeney, C.; Turk, M.; Höllerer, T. Multi-view gesture annotations in image-based 3D reconstructed scenes. In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, Munich, Germany, 2 November 2016; pp. 129–138.

108. Rice, M.; Chia, S.C.; Tay, H.H.; Wan, M.; Li, L.; Ng, J.; Lim, J.H. Exploring the Use of Visual Annotations in a Remote Assistance Platform. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, San Jose, California, USA, 7 May 2016; pp. 1295–1300.

109. Chang, Y.-C.; Wang, H.-C.; Chu, H.-K.; Lin, S.-Y.; Wang, S.-P. AlphaRead: Support Unambiguous Referencing in Remote Collaboration with Readable Object Annotation. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, Oregon, USA, 25 February 2017; pp. 2246–2259.

110. Tan, K.H.; Robinson, I.N.; Culbertson, B.; Apostolopoulos, J. Enabling genuine eye contact and accurate gaze in remote collaboration. In Proceedings of the 2010 IEEE International Conference on Multimedia and Expo, Singapore, 19-23 July 2010; pp. 430–435.

111. Gelb, D.; Subramanian, A.; Tan, K.H. Augmented reality for immersive remote collaboration. In Proceedings of the 2011 IEEE Workshop on Person-Oriented Vision 2011; pp. 1–6, doi:10.1109/pov.2011.5712368.

112. Huang, W.; Alem, L. Supporting Hand Gestures in Mobile Remote Collaboration: A Usability Evaluation. In Proceedings of the 25th BCS Conference on Human-Computer Interaction, Swindon, GBR, 4 July 2011; pp. 211–216.

113. Sodhi, R.S.; Jones, B.R.; Forsyth, D.; Bailey, B.P.; Maciocci, G. BeThere: 3D mobile collaboration with spatial input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April 2013; pp. 179–188.

114. Zenati-Henda, N.; Bellarbi, A.; Benbelkacem, S.; Belhocine, M. Augmented reality system based on hand gestures for remote maintenance. In Proceedings of the 2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, Morocco, 14–16 April 2014; pp. 5–8.

115. Gupta, K.; Lee, G.; Billinghurst, M. Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 2413–2422, doi:10.1109/tvcg.2016.2593778.

116. Higuch, K.; Yonetani, R.; Sato, Y. 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, San Jose, California, USA, 7 May 2016; pp. 5180–5190.

117. Li, J.; Manavalan, M.E.; D'Angelo, S.; Gergle, D. Designing Shared Gaze Awareness for Remote Collaboration. In Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion—CSCW '16 Companion, San Francisco, California, USA, 27 February 2016; pp. 325–328.

118. Lee, G.; Kim, S.; Lee, Y.; Dey, A.; Piumsomboon, T.; Norman, M.; Billinghurst, M. [POSTER] Mutually Shared Gaze in Augmented Video Conference. In Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Nantes, France, 9–13 October 2017; pp. 79–80.

119. Huang, W.; Billinghurst, M.; Alem, L.; Kim, S. HandsInTouch: Sharing gestures in remote collaboration. In Proceedings of the 30th Australian Conference on Computer-Human Interaction, Melbourne, Australia, 4 December 2018; pp. 396–400.

120. Otsuki, M.; Maruyama, K.; Kuzuoka, H.; Suzuki, Y. Effects of Enhanced Gaze Presentation on Gaze Leading in Remote Collaborative Physical Tasks. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, 21 April 2018; p. 368.

121. Shiro, K.; Okada, A.; Miyaki, T.; Rekimoto, J. OmniGaze: A Display-covered Omnidirectional Camerafor Conveying Remote User's Presence. In Proceedings of the 6th International Conference on Human-Agent Interaction, Southampton, United Kingdom, 4 December 2018; pp. 176–183.

122. Sasikumar, P.; Gao, L.; Bai, H.; Billinghurst, M. Wearable RemoteFusion: A Mixed Reality Remote Collaboration System with Local Eye Gaze and Remote Hand Gesture Sharing. In Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); Beijing, China, 10-18 October 2019; pp. 393–394.

123. Wang, P.; Zhang, S.; Bai, X.; Billinghurst, M.; He, W.; Wang, S.; Zhang, X.; Du, J.; Chen, Y. Head Pointer or Eye Gaze: Which Helps More in MR Remote Collaboration? In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; pp. 1219–1220.

124. Zakharkin, I.; Tsaturyan, A.; Cabrera, M.A.; Tirado, J.; Tsetserukou, D. ZoomTouch: Multi-User Remote Robot Control in Zoom by DNN-based Gesture Recognition. In Proceedings of the SIGGRAPH Asia 2020 Emerging Technologies, New York, NY, USA, 4 December 2020; pp. 1–2.

125. Xiao, C.; Huang, W.; Billinghurst, M. Usage and Effect of Eye Tracking in Remote Guidance. In *32nd Australian Conference on Human-Computer Interaction*, New York, NY, USA, 2 December 2020; pp. 622–628.

126. Teo, T.; Lee, G.A.; Billinghurst, M.; Adcock, M. Hand gestures and visual annotation in live 360 panorama-based mixed reality remote collaboration. In Proceedings of the 30th Australian Conference on Computer-Human Interaction, Melbourne Australia, 4 December 2018; pp. 406–410.

127. Kim, S.; Lee, G.; Huang, W.; Kim, H.; Woo, W.; Billinghurst, M. 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, Glasgow, Scotland, UK, 2 May 2019; p. 173.

128. Teo, T.; A Lee, G.; Billinghurst, M.; Adcock, M. Supporting Visual Annotation Cues in a Live 360 Panorama-based Mixed Reality Remote Collaboration. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23-27 March 2019; pp. 1187–1188.

129. Teo, T.; Lee, G.A.; Billinghurst, M.; Adcock, M. Investigating the use of Different Visual Cues to Improve Social Presence within a 360 Mixed Reality Remote Collaboration. In Proceedings of The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry, Brisbane, QLD, Australia, 14 November 2019; p. 4.

130. Bai, H.; Sasikumar, P.; Yang, J.; Billinghurst, M. 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, New York, NY, USA, 21 April 2020; pp. 1–13.