Eng. Sci., 2025, **37**, 1802

Engineered Science
DOI: https://dx.doi.org/10.30919/es1802

Extended Reality Remote Collaboration Supporting Visual Annotation Cues for Industry: A Literature Review

Peng Wang,[1,#] Yue Wang,[1] Yiwei Wang,[2] Mark Billinghurst,[3] Dongyu Yang,[4] Huan Yang,[5,#,*] Rong Luo[1] and Xiangyu Zhang[1]

Abstract

Extended Reality (XR) remote collaboration systems are crucial in advancing operational efficiency, adaptive capacity, and user satisfaction across industrial tasks within Industry 5.0. By using the unique capabilities of XR, remote collaborative systems enhance human adaptability and drive value-driven manufacturing processes, particularly in applications such as remote maintenance, training, and assembly. This review synthesizes insights from 69 peer-reviewed studies on XR remote collaboration that utilize visual annotation cues in industrial settings. It systematically analyzes display modalities, interaction methodologies, and shared visual cues, complemented by case studies of real-world implementations. The review critically evaluates existing literature, proposes improvements to current approaches, and explores alternative strategies. Importantly, this review offers a nuanced examination of the complexities surrounding XR-mediated collaborations, reinforcing the ongoing necessity and practical pathways for advancing industrial XR remote collaboration that integrate visual annotation cues.

Keywords: Extended reality (XR); Augmented reality (AR); Mixed reality (MR); Remote collaboration; Visual annotations cues; Industrial applications.
Received: 27 May 2025; Revised: 09 October 2025; Accepted: 15 October 2025
Article type: Review article.

1. Introduction

As the growing advancement of Extended Reality (XR) continues, innovative approaches such as remote collaboration[1,2] and co-located collaboration[3] are being developed to address the constraints of time and space in multi-user industrial sectors.[4] Remote collaboration refers to the process where geographically distributed individuals or teams utilize digital communication media to work together toward a shared objective. With the rapid development of technology, remote collaboration systems through XR – an umbrella term encompassing all real-and-virtual combined environments and human-machine interactions, including augmented reality (AR), virtual reality (VR), and mixed reality (MR) (see Fig. 1) – are progressing incrementally yet steadily, though they remain far from mainstream adoption.[1]

Additionally, XR has the three core principles: (i) 3D immersion, (ii) natural interaction, and (iii) spatial registration. In the vision of Industry 5.0, XR collaborative platforms position humans at the core of performing tasks, enabling real-time spatial interaction and shared decision-making in industrial processes.[5,6] XR remote collaboration systems can capitalize on human adaptability by integrating the strengths of XR, driving value-centric production in manufacturing, such as remote maintenance,[2,7] training,[8] assembly,[9–11] and design.[2,12]

In this review, we provide a thorough summary and analysis of the effect of visual annotation cues in XR remote collaboration for industry. It should be noted that we focused the synchronous remote collaboration. The significance of such systems has been underscored since the onset of the COVID-19 pandemic, alongside the current global geopolitical turmoil and conflicts.[13,14] Industrial distributed collaborators harness XR-driven spatial interaction and virtual-real fusion visualization to attain groundbreaking performance in operational tasks by supporting contextual visual annotation cues, thereby improving collaborative efficiency, situational awareness, user experience, and adaptive problem-solving capabilities.[1,7,12,15]

A cornerstone of XR remote collaboration lies in its

---

[1] *School of Integrated Circuits, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China*

[2] *School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China*

[3] *Auckland Bioengineering Institute, The University of Auckland, Auckland, 1010, New Zealand*

[4] *China Electronic Product Reliability and Environmental Testing Research Institute, Guangzhou, Guangdong, 510610, China*

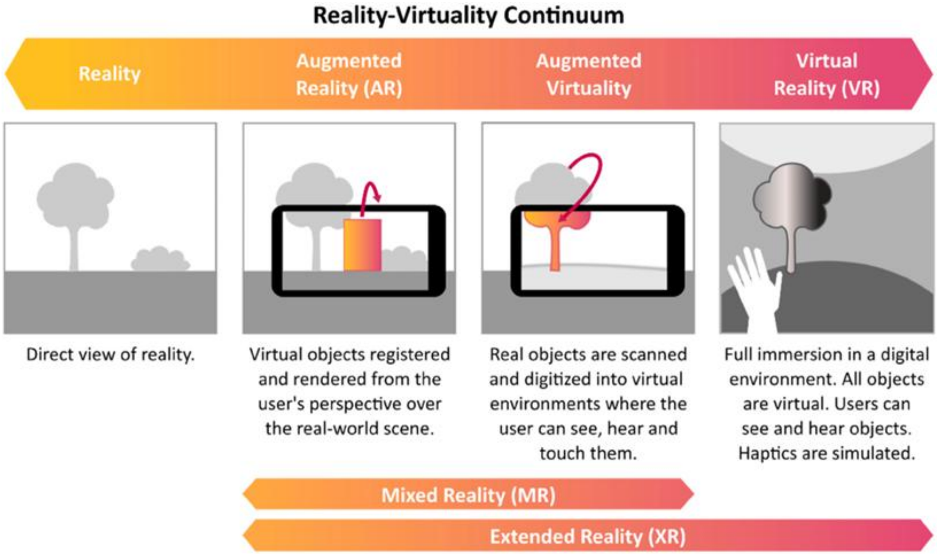


Fig. 1: XR as an umbrella term ranging from AR to VR. Reproduced from.[14]

ability to integrate visual annotation cues. In industrial maintenance scenarios, for example, XR systems enable technicians to conduct remote diagnostics and repairs by exchanging real-time contextual visual annotations cues, such as spatially anchored annotations and dynamic 3D virtual-real fusion. Local operators can create visual cues in real time to highlight position, while remote experts also provide step-by-step guidance instructions via superimposed annotations. This communication of visual information minimizes ambiguity, enhances task execution, and ensures alignment between on-site actions and expert instructions, thereby demonstrating their capacity to redefine industrial problem-solving.[7]

XR remote collaboration systems using visual annotation cues have emerged as a pivotal innovation in industrial applications, addressing critical challenges through three key advantages. First, visual annotations inherently enhance intuitive communication. Unlike text-based or verbal instructions, annotations enable local workers to pinpoint parts, components, or highlight potential anomaly locations. More importantly, users can make annotations by gesture-based interaction or controller-based interaction in a natural and intuitive way. This mimics real-world annotations-making, such as circling defects or sketching solutions directly in XR collaborative environments. Second, the systems provide spatial awareness, as annotations are anchored to real-world coordinates and can retain their positional relevance, ensuring that all collaborators perceive and interact with digital cues in a shared 3D XR workspace. Finally, visual annotations significantly improve operational efficiency. Real-time annotations allow remote collaborators to sketch repair plans, flag critical areas, or demonstrate procedures spatially, transforming abstract concepts into tangible, step-by-step actions. Such clarity minimizes training time, accelerates task completion, and reduces errors caused by ambiguous guidance.

The innovation of this work lies in its dual focus. First, this is the first comprehensive review (2019-2024) specifically examining how visual annotation cues shape collaboration effectiveness in industrial XR environments, compared to the related works.[2,6,13,15-19] Second, three doctoral research projects[20–22] focused on XR remote collaboration from 3D/360° capture and guidance cues. While previous works have touched on related aspects, this review advances the field by exploring XR remote collaboration supporting visual annotations cues for industry. This novel work not only extends existing scholarship but also provides actionable insights for optimizing XR remote collaborative tools in alignment with human-centric manufacturing.

2. Related works

The growing adoption of visual annotation cues in industrial XR remote collaboration highlights their transformative potential as a rapidly evolving domain, driving advanced investment from both academia and industry. While existing reviews have explored AR/MR/XR collaborative systems, specifically, XR assembly and maintenance.[6,13,15] XR remote human-robot interaction,[16] AR/MR remote collaboration,[2,17] and VR/AR in manufacturing (*i.e.*, design, assembly, training, repair),[18,19] as shown in Fig. 2. Therefore, we find a critical gap persists in exploring the effect of visual annotation cues in XR remote collaboration for industry. Bridging this knowledge gap is essential to unlock the full potential of XR remote collaboration, where visual annotations cues act as the communication bridge between distributed collaborators, reducing errors, improving decision-making, and aligning with Industry 5.0's vision of human-centric, adaptive manufacturing systems. Therefore, we explore a comprehensive review of the effect of visual annotation cues in XR remote collaboration for industrial applications from 2019 to 2024, extending previous studies.[2,12,15]

---

[5] *Big Data Institute, Baoshan University, Baoshan, Yunnan, 678000, China*

[#] *These authors contributed equally to this work*

[*] *Email:* d220201037@stu.cqupt.edu.cn (Huan Yang)

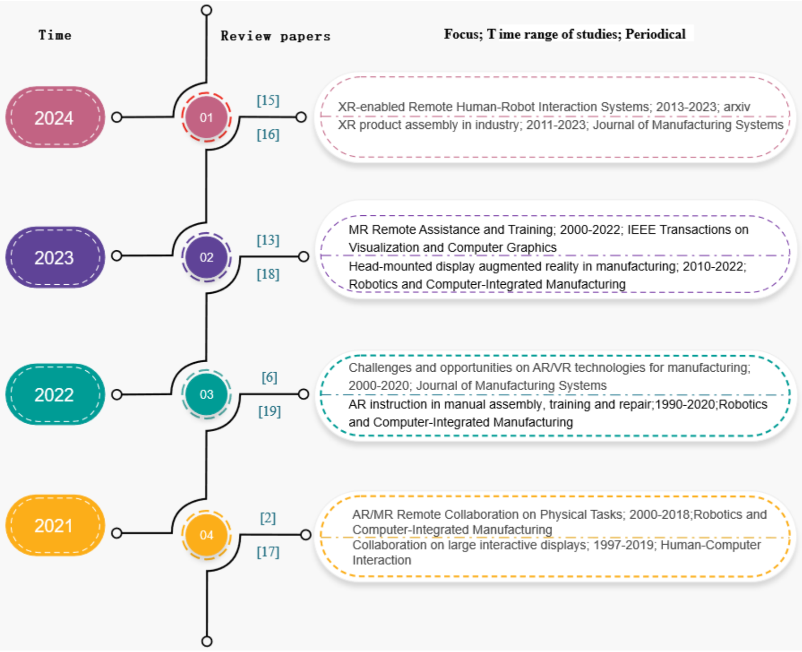


Fig. 2: Related reviews. Reproduced from.[2,6,13,15–19]

Moreover, our search revealed three doctoral theses[20–22] that align closely with the scope of this review. Specifically, Theophilus Teo's research[22] proposes a hybrid MR system combining 360° video and 3D reconstruction to enhance remote collaboration, demonstrating advantages for complex tasks requiring detailed object inspection and environmental navigation. While hybrid systems showed superior capability for specific challenges, users noted steeper learning curves and perceived simpler systems as adequate for basic tasks. Bernardo Marques' thesis[21] explores AR's potential in remote collaboration, addressing evaluation challenges through a human-centered framework, and proposes the CAPTURE toolkit for contextualized data analysis, demonstrating effectiveness in AR maintenance tasks and broader collaborative scenarios. Lei Gao's thesis[20] explores MR room-scale remote collaboration, developing AR/VR prototypes with 3D capture and guidance cues. Studies found that 3D environments boost experts' spatial awareness, high-resolution views aid complex tasks, and mutual awareness improves communication, informing interface design principles and future directions.

This literature review is motivated by the imperative to evaluate the transformational capacity of XR in enabling remote collaboration, particularly through XR collaborative systems using visual annotation cues, within industrial cooperation scenarios. Focused on distributed industrial collaboration, the investigation pursues three primary objectives: first, to conduct a systematic evaluation of proven advantages of XR interfaces using visual annotations in improving manufacturing processes; second, to investigate technical implementation methodologies from interaction technique and XR displays; and third, to deliver practical XR development frameworks and specialized hardware solutions for building annotation-enhanced remote collaboration platforms. Through this comprehensive analysis, the review aims to equip both academic researchers and industrial stakeholders with evidence-based strategies for creating XR remote collaborative systems for decentralized production ecosystems. The subsequent sections present the paper's structured organization. Section 2 presents related works, and Section 3 provides the research background and review method. Section 4 presents a systematic scoping survey. Then Section 5 reports on the discussion. Next, we provide the future directions in Section 6. Finally, we concluded the review.

3. Background and review method

3.1 XR remote collaboration using visual annotations cues in industry

XR remote collaboration using visual annotation cues in industrial settings, using gesture-based interaction and other human-computer interaction (HCI) methods to connect geographically dispersed users within shared XR environments. By employing visual spatial annotations and contextual XR visualizations, these systems facilitate real-time shared spatial awareness and immersive engagement enhancing industrial tasks such as maintenance, assembly, and operational guidance, and design.[2,6,7,15]

3.2 Visual annotations cues

In XR remote collaboration, visual cues(*e.g.*, gestures,[11,23,24] eye gaze or head pointer,[25] and annotations[10,24]) are essential for conveying information between geographically dispersed collaborators. Specifically, visual annotation cues such as sketches serve as a powerful tool for efficiently conveying ideas and guidance. These cues can transcend language barriers, enabling collaborators to communicate complex concepts and spatial relationships with ease. Unlike text or

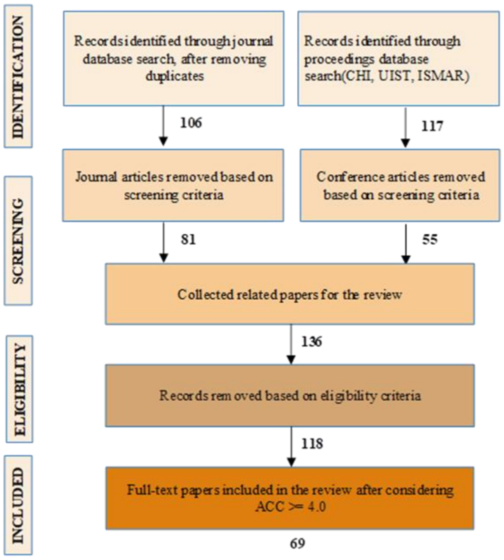


Fig. 3: PRISMA flowchart of the paper collection process.

Table 1: Keywords and search string.

	AND				OR
OR	eXtended Reality(XR)	Remote	Collaboration	Annotations	Assembly
	Mixed Reality (MR)	Tele-	Collaborative	Sketches	Maintenance
	Augmented Reality (AR)	Distributed	Communication	Visual Cues	Design
	Virtual Reality (VR)			Cues	Assistance
					Mentoring

verbal descriptions, visual annotations in XR remote collaboration allow collaborators to engage in industrial tasks in real-time.[10,11,26] Importantly, natural gesture-based interactions are highly effective for quickly providing context-aware annotations, which can be spatially anchored within 3D XR environments, enhancing the intuitiveness and efficiency of collaborative tasks.

3.3 PRISMA method

In performing this structured analysis, we adhered to established PRISMA protocols for academic surveys.[27–29] We conducted a search for pertinent research and evaluated its eligibility based on inclusion and exclusion criteria. Specifically, articles were included if they focused on XR remote collaboration using visual annotation cues in industrial settings, published from 2019 to 2024, and written in English. Conversely, articles were excluded with full-text unavailable and a few pages. This screening process followed the methodology outlined in Fig. 3.

We identified pertinent articles from prominent journals and conferences from Web of Science, Elsevier, Springer database for journal articles, ACM Digital Library, and IEEE Explore for conference papers (*e.g.*, the CHI Conference on Human Factors in Computing Systems (CHI), IEEE International Symposium on Mixed and Augmented Reality (ISMAR)). The search was conducted in March 2025, and included articles in the English language from the past six years (2019-2024). The search was conducted by combining the terms outlined in Table 1. Using these keywords, we searched the field topic, resulting in an initial total of 223 articles after duplicate removal. Following the application of screening criteria, 136 articles were considered eligible and selected for further analysis.

To assess the academic influence of the identified publications, we utilized the Average Citation Count (ACC), computed as ACC = Total citations / Publication age (in years).[30,31] Citation data were gathered from Google Scholar on February 14, 2025. Our selection criteria emphasize survey literature with an ACC of 4.0 or higher, ensuring the inclusion of works with lasting scholarly significance in the field. Notably, publications from 2024 were excluded from ACC calculations due to their recent publication date. Ultimately, a total of 69 papers were collected.

3.4 Research questions

To perform a comprehensive analysis of the effect of visual annotations cues in XR remote collaboration in industry, we devised the following four research questions:

(1) What is the current state of areas for XR remote collaboration using visual annotations cues in industry? (Section 3.1, 3.2, 3.3, and 3.4)

(2) What is the effect of supporting visual annotation cues

Table 1: XR remote collaborative systems supporting visual annotation cues. ♥= The system supports this feature.

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2024	[10]	♥	♥	♥	♥			GDI		♥		GDI	Providing Assistance/ Assembly	HTC VIVE Pro 2, MRTK, Vuforia Engine, HoloLens, Leap Motion	China	4
2024	[32]	♥	♥	♥	♥			CBI			HHD	GDI	Assembly	Mate Quest2, ARKit, Webkit, WebSocket API[1] WebRTC API[1],WebXR[1], 3D scanner app[1]	Australia, USA, South Korea	3
2024	[33]		♥				HHD	GDI			HHD	GDI	Providing Assistance	Vuforia, Photon Unity Networking[3]	Canada	0
2024	[34]		♥				HHD ♥	KBM, GDI		♥	HHD	GDI	MaintenanceTraining	HoloLens 2, WebRTC, Vuforia	Portugal	0
2024	[35]		♥		♥			GDI	♥			GDI	Discussion	Oculus, Meta Avatars SDK, Photon Engine PUN2	Australia, Denmark	12
2024	[36]		♥		♥			CBI	♥			CBI	Discussion	Meta Quest Pro	Australia Denmark	5
2024	[37]		♥		♥			CBI	♥			CBI	Discussion	Meta Quest Pro	Australia Denmark	4
2024	[38]	♥		♥	♥			CBI	♥			CBI	Assembly	Oculus Quest 2	Netherlands, Japan	5
2024	[39]	♥	♥				♥	KBM			♥	KBM	Providing	---	France	1

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2024	[25]		♥	♥EG	♥			GDI		♥		GDI	Assistance Providing Assistance/ Assembly	AR Meta2, FOVE, Ricoh Theta V, Leap Motion	South Korea Australia	0
2024	[40]		♥				HHD	GDI			HHD	GDI	Maintenance	WebRTC	Denmark, Australia	1
2024	[41]	♥	♥		♥			CBI	♥			CBI	Training	---	USA	0
2024	[42]	♥	♥	♥	♥			CBI		♥		GDI	Teaching	HoloLens 2, Meta Oculus Quest 2, Insta360 X3, Instant-NGP[5], Segment-Anything[1]	Canada	4
2024	[26]	♥	♥	♥	♥			CBI	♥			CBI	Training	VRTK, MeVisLab, HTC Vive Pro, Avatar SDK	USA Germany	11
2024	[43]	♥	♥				♥	KBM			HHD	GDI	Inspection	QR code	Canada	1
2024	[44]	♥	♥				♥	KBM		♥		GDI	Assembly	Dynamics guides 365 software, hololens 2	Australia	3
2024	[45]	♥	♥		♥			CBI			HHD	GDI	Discussion	PUN 2, Agora Software-Defined Real-time Network, RICHO THETA Z, HTC	Japan	0

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2024	[11]	♥	♥	♥	♥			GDI		♥		GDI	Assembly	VIVE Pro MRTK, HTC VIVE Pro 2, Leap Motion, Intel® RealSense D435i, HoloLens, WampServer	China	1
2023	[24]	♥	♥	♥	♥			GDI		♥		GDI	Providing Assistance/ Assembly	HTC VIVE Pro 2, MRTK, Vuforia Engine, HoloLens, Leap Motion, DBSCAN algorithm[1]	China	20
2023	[46]	♥	♥	♥	♥			GDI		♥		GDI	Providing Assistance/ Assembly	Point Cloud Library (PCL), HTC VIVE Pro 2, MRTK, Intel's RealSense SDK 2.0, Vuforia Engine, HoloLens, Leap Motion, Intel®	China	34

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2023	[47]	♥	♥	♥	♥			GDI		♥		GDI	Assembly	RealSense D435i MRTK, HTC VIVE Pro 2, Leap Motion, Intel® RealSense D435i, HoloLens, WampServer	China	6
2023	[48]		♥				♥	GDI			♥	GDI	Training	AR toolkit, SteamVR, Vuforia	China, Saudi Arabia	10
2023	[49]		♥			♥		GDI			♥	GDI	Assembly	HoloLens2, Photon Engine[1], Socket.io library[1]	USA	8
2023	[50]		♥		♥			CBI		♥		GDI	Providing Assistance	HoloLens2, HTC Vive Pro Eye, Logitech Ink Pen , Azure Kinect, Vuforia	Germany	8
2023	[51]		♥				♥	KBM			♥	KBM	Assembly	Azure Kinect camera , React framework[1], React Conva[1] , AR tags,	USA, Switzerland	4

Year	Ref	Visual cues			Remote site				Local site							
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques	Tasks	XR Devices/ Toolkits	Region	ACC
2023	[52]		♥				♥	KBM			♥	KBM	Assembly	workbench kit[1] Azure Kinect camera, React framework, AR tags, workbench kit	USA, Switzerland	4
2023	[53]		♥		♥			GDI	♥			GDI	Design	Oculus XR Interaction Toolkit, Meta Avatars SDK, Meta Quest 2	Denmark, Australia	44
2023	[54]	♥	♥	♥	♥			CBI		♥		GDI	Assembly	Oculus Quest, HoloLens 2, Vuforia Spatial Toolbox Virtualizer[3], Azure Kinect	Australia	32
2023	[55]		♥				♥	KBM			♥	KBM	Discussion	---	Austria	8
2023	[56]		♥	♥	♥			GDI		♥		GDI	Assembly	Meta2 AR HMD, Ricoh Theta V, Leap Motion, FOVE VR HMD	South Korea, Australia	7
2023	[23]	♥	♥	♥	♥			GDI		♥		GDI	Assembly	PCL, MRTK, HTC VIVE PRO2,	China, New Zealand	6

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2023	[57]		♥		♥			CBI		♥		GDI	Discussion	Hololens 2, YOLOv5, Leap Motion WebRTC, Ready Player Me[1], Headshot plugin[1], 360 camera (Ricoh Theta), HTC VIVE Pro, Meta Quest 2	USA	10
2023	[58]		♥				HHD	GDI		♥		GDI	Providing Assistance	HoloLens2, Tablet, Dynamics 365 Remote Assist and Teams, DemoPro[1]	France, USA	9
2023	[59]	♥	♥				♥	KBM			HHD	GDI	Inspection	Oculus Quest, Zoom and Microsoft Teams	Canada, New Zealand, Australia, Chile	11
2023	[60]		♥				♥	KBM			HHD	GDI	MaintenanceTraining	WebRTC, Vuforia	Portugal	14
2022	[61]		♥				♥	KBM		♥		GDI	Assembly design	Intel RealSense D435i, Zbar and OpenCV, WebRTC	China	5

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
														, PCL, Hololens 2, Vuforia Chalk[1], Scope AR[1]		
2022	[62]		♥				♥	KBM		♥		GDI	Providing Assistance	Dynamics 365 Remote Assist, Hololens	USA	4
2022	[63]		♥	♥EG	♥			GDI		♥		GDI	Discussion	Ricoh Theta V 360° , HoloLens2, HTC Vive Pro Eye, Leap Motion, SRanipal SDK, MRTK, Ultraleap Gemini libraries[1], Vuforia	Australia	9
2022	[64]		♥					KBM			HHD	GDI	Inspection	SLAM, 2D markers, Tesseract-OCR[2]	South Korea	5
2022	[65]		♥	♥EG	♥			GDI	SAR			KBM	Assembly	OpenCV, aGlass-DKII, Leap Motion, HTC Vive	China	11
2022	[66]	♥	♥				♥	KBM			HHD	GDI	Inspection	ARCore SDK	Iran, USA	24
2022	[67]		♥				♥	KBM			HHD	GDI	Maintenance Training	WebRTC, Vuforia	Portugal	16.5

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2022	[68]		♥				♥	KBM			HHD	GDI	Maintenance Training	WebRTC, Vuforia	Portugal	35
2021	[69]		♥				♥	KBM		♥		GDI	Maintenance Training	Microsoft HoloLens 2, WebRTC, Vuforia	Portugal	6
2021	[70]		♥	Image			♥	KBM			HHD	GDI	Providing Assistance	ARcore, Twilio[2]	Italy	8
2021	[71]	♥	♥		♥			CBI			HHD	GDI	Assembly	Vufroia, VR HMD	South Korea	22
2021	[72]		♥				♥	KBM		♥		GDI	Assembly	HoloLens	Germany	6
2021	[73]		♥		♥			CBI	♥			CBI	Design	Photon Unity Networking, HTC Vive	Thailand	22
2021	[74]		♥		♥			CBI		♥		GDI	Providing Assistance	HTC Vive Pro, IMU-equipped gloves[2], Hololens ,Vuforia, WebRTC	USA	44
2021	[75]		♥		♥			CBI		♥	HHD	GDI	Providing Assistance	HTC VIVE Pro Eye, Tobii Eye Tracking , Unity UNet networking, HoloLens 2, Azure Kinect camera with the Microsof	Germany	25

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
														t Body Tracking SDK		
2021	[76]		♥				♥	KBM				GDI	Providing Assistance	Zoom	USA	12.7
2021	[77]		♥		♥			CBI	♥		HHD	CBI	Design	VRChat SDK[2], Oculus Rift, Udon 3D pen[2]	USA	25.7
2021	[78]		♥				♥HHD	KBM, GDI				GDI	Maintenance Training	WebRTC, Vuforia	Portugal	12
2020	[79]		♥				♥	KBM		♥		GDI	Maintenance	WebRTC, WebSocket, SLAM, Halo Mini holographic	China	10.5
2020	[80]		♥	♥	♥			GDI	SAR			KBM	Assembly	HTC Vive, Leap Motion, Leap Motion SDK	China, Australia	15.5
2020	[81]	♥	♥				♥	KBM		♥		GDI	Maintenance	HoloLens	United Kingdom, Slovenia	9.2
2020	[82]		♥	♥	♥			CBI		♥		GDI	Assembly	Leap motion, Meta2	Australia	12
2020	[83]		♥	EG	♥			CBI	SAR		HHD	KBM	Assembly	HTC Vive Eye Pro, aGlass, Projector, Opencv	China, Australia	9.5
2020	[84]		♥				HHD	GDI				GDI	Maintenance	ARCore	Australia	11.5
2020	[85]	♥	♥				♥	KBM		♥		GDI	Maintenance	HoloLens, MRTK,	Greece	43

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2019	[86]		♥	♥	♥	♥		GDI	♥	♥		GDI	Design	Vuforia, Unity UNet API Kinect, HTC Vive, ZED Mini[1]	Canada, USA	30.4
2019	[87]		♥		♥			CBI	♥			CBI	Discussion	HTC Vive, HTC Vive Pro	Australia, New Zealand	30
2019	[88]		♥	♥			HHD	GDI		♥		GDI	Assembly	OpenCV, Logitech H230, Vuzix Wrap 1200	Australia	12.8
2019	[89]		♥	♥	♥			GDI		♥		GDI	Assembly	Meta2, FOVE VR HMD, a Leap Motion	Australia, South Korea	29.6
2019	[90]		♥	♥HP	♥			CBI	SAR			KBM	Assembly	Opencv, HTC Vive, a Leap Motion	China, Australia	8
2019	[91]		♥	♥HP	♥			CBI	SAR			KBM	Assembly	Opencv, HTC Vive, a Leap Motion	China, Australia	14.4
2019	[92]		♥				♥	KBM			HHD	GDI	Design	Vuforia, Google Project Tango Development Kit	Italy	14.8
2019	[93]		♥				♥	KBM			HHD	GDI	Inspection	Vuforia, ARKit	Spain	12

Year	Ref	Visual cues			Remote site				Local site				Tasks	XR Devices/ Toolkits	Region	ACC
					HMD				HMD							
		VRP	Annotations	Gestures	VR	MR	Desktop	Interaction techniques	VR	MR	Desktop	Interaction techniques				
2019	[94]	♥	♥			♥		KBM			HHD	GDI	Inspection Maintenance	Picking[1], hybrid tracking method[1], Oculus Rift, FLIR ONE camera, Vuforia, Connectify[1]	Canada	33.6

(3) What devices/toolkits are used in developing platforms? (Section 3.8)

(4) What are the main future trends of XR remote collaboration using visual annotation cues? (Section 5)

These research questions directed our literature review, organized into focused sections, to enhance the integration of visual annotations cues in industrial XR remote collaboration and offer practical insights for future research and real-world applications.

4. Systematic scoping survey

4.1 Annual publication trends

We systematically analysed the collected 69 papers from visual cues, XR display type, and interaction techniques for remote/local sites, collaborative tasks, XR devices/toolkits, region, and ACC as shown in Table 2. Fig. 4 shows the number of papers on XR remote collaboration supporting visual annotation cues. The overall trend has been rising year by year in the case of the value of ACC. This indicates a growing interest in XR remote collaboration supporting visual annotation cues in industry.

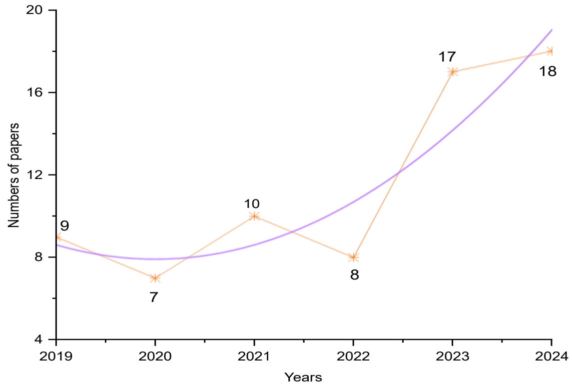


Fig. 4: The published papers numbers with years.

4.2 Document type

We analyzed the publication sources of the 69 collected papers and found that 42 were journal articles and 27 were conference papers. The statistics on the number of papers published in different journals and conferences are shown in Fig. 5 and Fig. 6. Fig. 5 shows that all journals of publications, such as "International Journal of Advanced Manufacturing Technology (IJAMT)", "Advanced Engineering Informatics", "IEEE Transactions on Visualization and Computer Graphics", etc. In addition, Fig. 6 shows that all conferences of publications, CHI, ISMAR, UIST, etc. The higher proportion of journal articles suggests the surveyed domain prioritizes in-depth technical validation and longitudinal studies, which are hallmarks of mature research areas. Conference venues (*e.g.*, CHI, ISMAR, UIST), while prestigious, often feature cutting-edge but preliminary work; the journal majority implies consolidation of established knowledge. In addition, the prominence of top-tier conferences (CHI, UIST) highlights that the conference track remains vital for high-impact innovation, particularly in human-centered domains (*e.g.*, XR, UI/UX). Their inclusion signifies that emerging trends and prototypes are actively debated within these communities before maturing into journal articles.

4.3 XR Remote collaboration using visual annotation cues in industry

Based on Wang *et al.*,[15] and industrial application scenarios, XR remote collaboration using visual annotation cues can be categorized into three sequential stages: before assembly, on-site assembly, and after assembly, as shown in Fig. 7. These applications are deployed across various hardware platforms From the perspective of XR remote collaboration using visual annotation cues in industrial settings, this approach unlocks transformative advantages by improving natural, intuitive

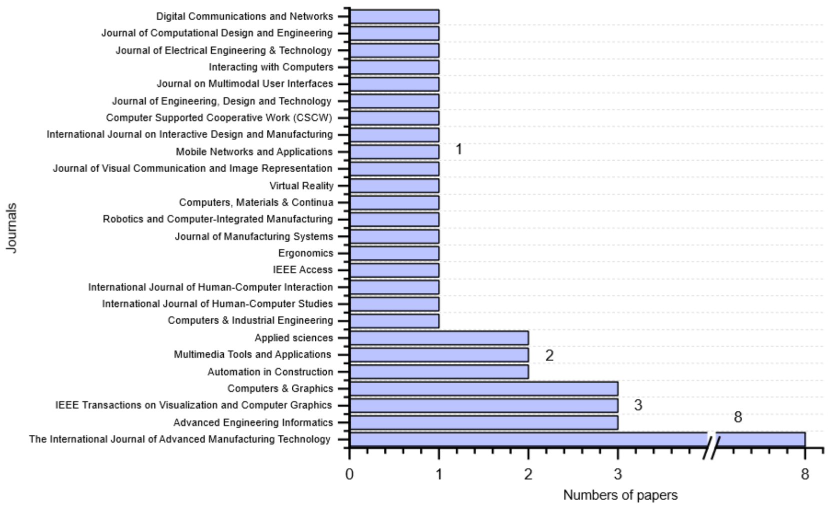


Fig. 5: Publication of journal articles.

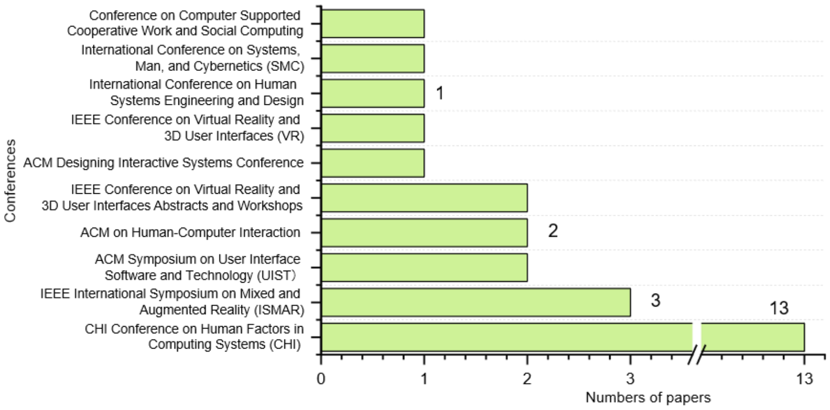


Fig. 6: Publication of conference articles.

interactions through contextual annotations, dynamic labels, and 3D-guided instructions.[21] These visual annotations cues enhance workforce training agility, elevate operational performance, and improve user interaction experience by enabling real-time collaborative problem-solving, immersive skill transfer, and assembly/maintenance guidance across distributed industrial ecosystems.[21,32,34,38]

From Table 3, we categorized XR remote collaborative applications using visual annotations cues in industry as illustrated in Table 3. The relevant studies of "before assembly" 21] These visual annotations cues enhance workforce training agility, elevate operational performance, and improve user interaction experience by enabling real-time collaborative problem-solving, immersive skill transfer, and assembly/maintenance guidance across distributed industrial ecosystems.[21,32,34,38]

From Table 3, we categorized XR remote collaborative applications using visual annotations cues in industry as illustrated in Table 3. The relevant studies of "before assembly" (18.7%, 14/75) include discussion and design. The field of "on-site assembly" (41.3%, 31/75) includes collaborative

4.4 Region

From Table 2, we listed the authors' countries. Fig. 8 illustrates the distribution of data across five regions. More specifically, this pie shows that South America (26%) and Asia (24%) hold similarly high proportions, followed by North America (27%) and Europe (22%). Oceania (1%) accounts for the smallest share. The near-equal distribution across different regions, excluding Oceania, reflects balanced yet context-driven global engagement in industrial XR remote collaboration research. More specifically, South America's strong representation correlates with manufacturing digitization in countries where XR supports remote expert guidance for machinery maintenance and workforce training in geographically dispersed facilities. Asia's significant contribution aligns with leadership in industrial automation (*e.g.*, China, Japan, South Korea) and smart factory initiatives, where visual annotation systems enhance precision manufacturing and supply chain coordination. North America and Europe focus on high-fidelity industrial workflows using XR, evidenced by Boeing's use of real-time 3D model collaboration and Airbus' reported 6-month reduction in aircraft development cycles. Conversely, Oceania's minimal output (1%) highlights structural constraints: limited heavy manufacturing sectors, smaller R&D funding pools prioritizing mining/agriculture over industrial XR, and geographic isolation reducing multinational trial opportunities.

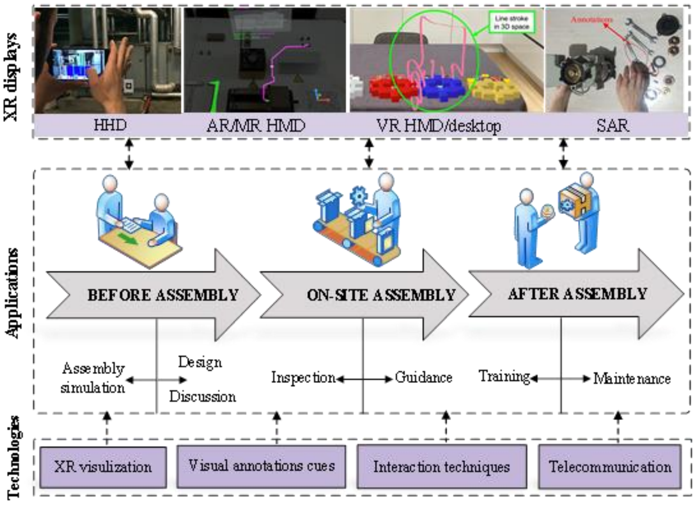
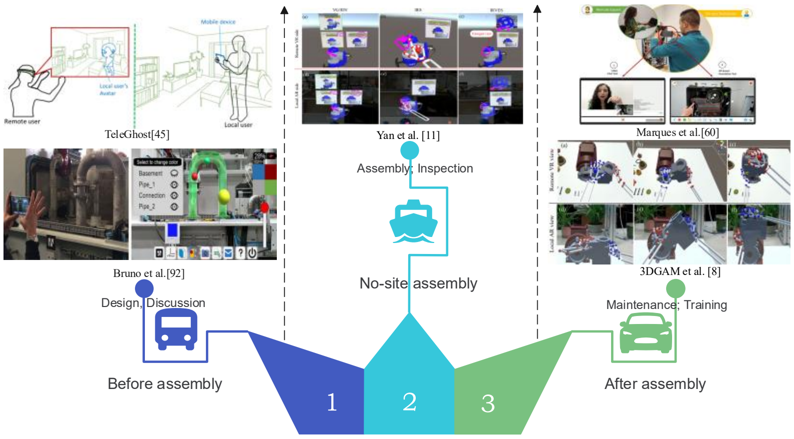


Fig. 7: Industrial XR remote collaboration using visual annotation cues.

Table 3: XR remote collaborative applications supporting visual annotations cues in industry.

	Before assembly	No-site assembly	After assembly
Applications			
Refs	[35,37,45,53,55,57,63,73,77,86,87,92,95]	[10,11,23–25,38,43,44,46,47,49,51,52,54,56,59,64–66,71,72,80,83,88–94]	[10,24–26,32–34,39–42,46,48,50,58,60,62,67–70,74–76,78,79,81,84,85,94]
Count	14	31	30

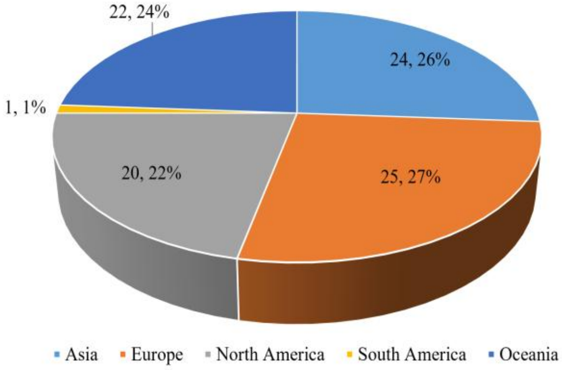
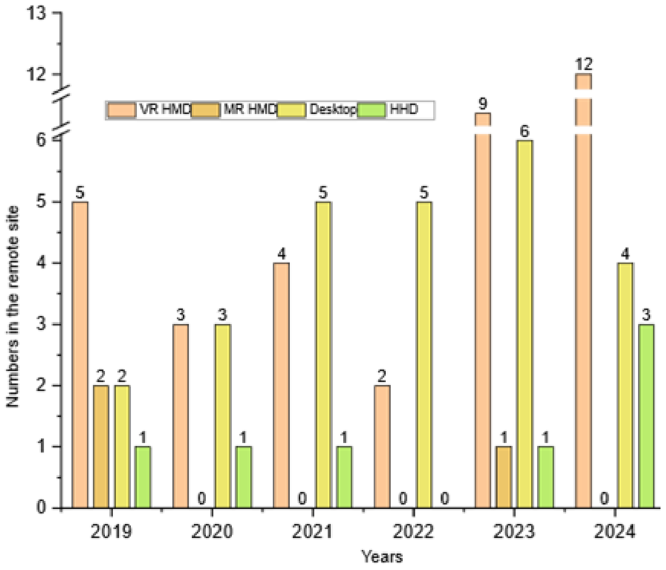


Fig. 8: Regional distribution.

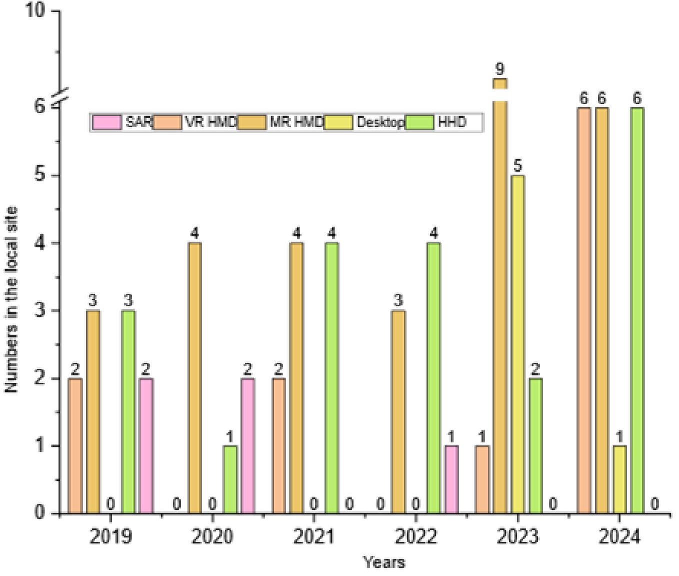
4.5 Display mode

Based on the findings presented in Table 2, we concluded that XR displays mode primarily encompasses HMDs(*i.e.*, VR and MR), (SAR), desktop-integrated visualization systems, and HHDs, as illustrated in Fig. 7. A closer examination of these technologies reveals distinct trends in their adoption, particularly in remote sites. Specifically, VR HMDs emerge as the dominant display technology in remote environments, and the overall distribution of XR visualization is depicted in Fig. 9(a). Furthermore, Fig. 10(a) provides a detailed breakdown of usage, showing that VR HMDs account for 49% of the total, followed by desktop systems at 37%, HHDs at 10%, and SAR constituting merely 4% of the usage. This distribution underscores a marked preference for immersive VR HMDs over traditional setups, highlighting the growing reliance on highly immersive technologies in XR visualization applications.

For the local site, Fig. 9(b) shows the overall distribution of XR visualization. In addition, Fig. 10(b) illustrates the proportion of XR visualization, showing that MR HMDs account for 40% of the total, followed by HHDs at 27%, VR HMDs at 15%, SAR at 10%, and the desktop system at merely 8% of the usage. The dominance of MR HMDs (40%) reflects their pivotal role in bridging physical and digital workspaces, which is essential for industrial visual annotation tasks. This aligns with market trends, where MR headsets are becoming the primary hardware investment, surpassing traditional VR devices due to their ability to overlay contextual cues directly onto machinery and workspaces. HHDs rank second, highlighting their persistent utility for quick inspections, mobility in large facilities, and lower entry barriers—especially in field services where bulkier headsets are impractical. VR HMDs (15%) maintain relevance primarily in training simulations and design reviews, where full immersion benefits complex spatial understanding, though their detachment



(a)XR visualization in the remote site



(b)XR visualization in the local site

Fig. 9: Trend of XR visualization over time.

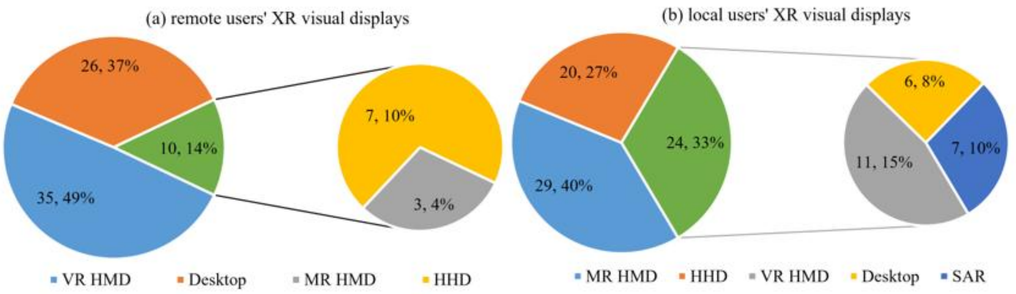


Fig. 10: The proportion of XR visualization, (a) XR visualization in the remote site, (b) XR visualization in the local remote site.

from physical environments limits real-time annotation utility. SAR—such as projectors or displays embedded in environments—serves niche applications like collaborative assembly, where shared visualizations across multiple users are critical. Conversely, desktop systems (8%) trail significantly, as their non-immersive nature fails to support intuitive spatial annotation or hands-free operation, which are essential for industrial workflows like remote maintenance.

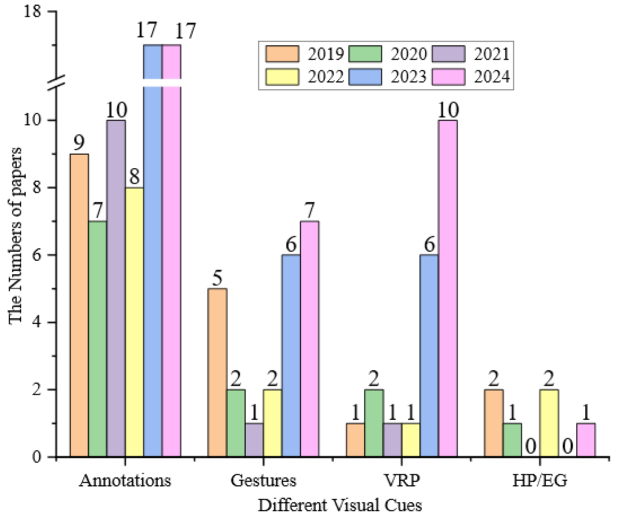


Fig. 11: The trend of providing visual cues.

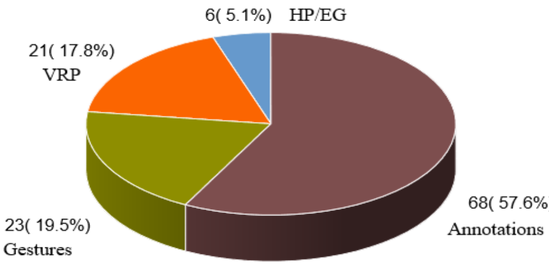


Fig. 12: The proportion of providing visual cues.

4.6 Visual information

In asymmetric remote collaborations, where a remote expert guides a local worker, visual annotation cues excel at conveying expertise into spatially intuitive cues. More importantly, gesture-based interaction plays a pivotal role in creating intuitive visual annotation cues within XR remote collaboration for industrial activities. By enabling collaborators to dynamically annotate the shared XR workspace, highlight critical components, or simulate physical actions (*e.g.*, pointing to a tool or mimicking a rotation), gestures align seamlessly with human-centric communication patterns. These gesture-driven annotations bridge the gap between abstract instructions and actionable guidance, particularly in hands-on tasks such as equipment assembly, maintenance, or repair. Furthermore, combining visual annotations with VRP or gestures/HP/EG allows collaborators to layer procedural guidance (*e.g.*, step-by-step markings) directly onto XR workspaces.[1,2,13]

As outlined in Table 2, the primary visual instruction modalities in XR remote collaboration for industrial applications comprise annotations, gestures, VRP, and HP/EG. Fig. 11 depicts the annual publication trends across these four categories from 2019 to 2024, while Fig. 12 highlights the proportional distribution of research focus among visual instruction types. Notably, visual annotations cues constitute the predominant category, representing 57.6% of visual cues, followed by VRP (19.5%), gestures (17.8%), and HP/EG cues (5.1%). Crucially, the multimodal integration of annotations with gestures or VRP demonstrates notable potential for advancing spatially contextualized guidance in industrial XR collaboration.

Based on the comprehensive analysis, we found that in the context of XR remote collaboration for industry, highlights and freehand drawing emerge as the most effective visual annotation cues, particularly when balancing clarity, flexibility, and user engagement.[21] More specifically, highlights are highly effective because they provide immediate visual salience without overwhelming the user. In industrial tasks where attention must be drawn to specific components or areas—such as identifying a faulty part or a step in an assembly sequence—highlights (*e.g.*, color overlays or glowing outlines) allow remote experts to quickly guide local workers' attention without cluttering the visual field. This is especially useful in environments with a limited field of view, such as HoloLens 2, where spatial precision and minimal distraction are critical.

Freehand drawing, on the other hand, offers expressive flexibility that static symbols like arrows or bounding boxes cannot match. It allows remote experts to sketch instructions directly onto the 3D space—such as circling a bolt to be tightened or drawing a path for cable routing—which provides natural human communication and supports nuanced, context-specific guidance. This method is particularly valuable in unstructured or unfamiliar tasks where predefined cues fall short. Moreover, freehand annotations can persist spatially, enabling asynchronous review and reducing cognitive load during real-time collaboration.

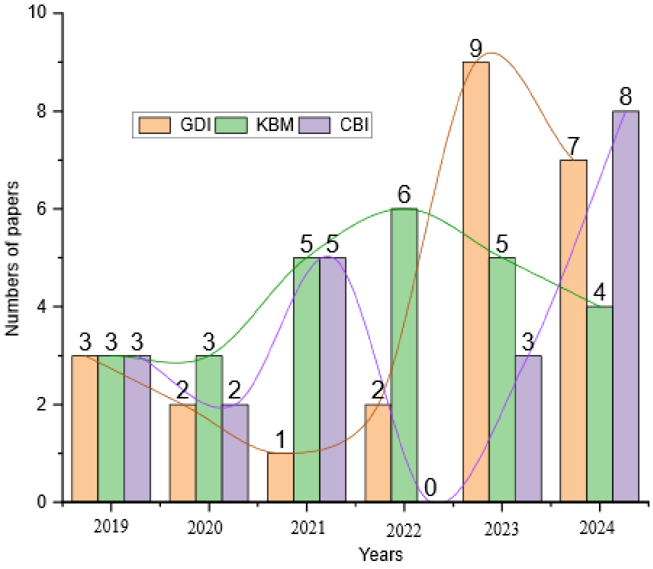
While other cues like 3D pointers or arrows are useful for directing attention or indicating direction, they lack the semantic richness and task-specific adaptability that highlight and freehand drawing provide. In sum, the combination of highlights for clarity and freehand drawing for expressiveness offers the most effective and human-centric approach to visual annotation in XR industrial collaboration.

In addition, after reviewing the literature, we conclude that there is currently no universally adopted semantic standard for the meaning of annotation cues in XR remote-collaboration systems used in industry. 1) Color "red" is not consistently mapped to "danger." In the empirical studies we surveyed, red is most often employed simply as a high-salience highlight to attract attention or to stress a step, rather than to encode hazard level. 2) Arrows, similarly, lack a fixed semantics. While many authors use arrows to indicate "next action" (*e.g.*, "insert here," "turn clockwise"), this is a design convention, not a normative rule. Other papers use arrows merely to point at regions of interest, leaving the intended action to be conveyed verbally or textually.

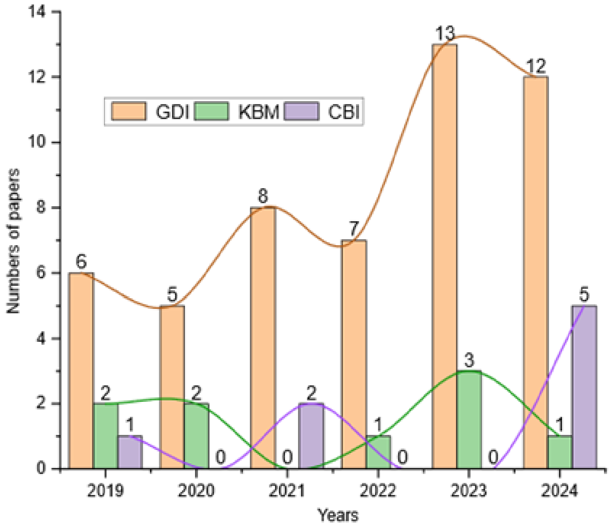
4.7 HCI techniques

Prior research in industrial XR-enabled remote collaboration has revealed distinct contextual challenges faced by distributed teams operating in XR environments, particularly regarding the effective integration of visual annotation systems with HCI paradigms.[2,6,13] As evidenced in Table 2, the systematic analysis of interaction techniques reveals three primary HCI approaches: gesture-driven interaction (GDI), controller-based interaction (CBI), and traditional keyboard-mouse interaction (KBM) for industrial asset manipulation and spatial manipulation of annotations.

Fig. 13 shows the trend of HCI methods for remote and local sites over time. There are no obvious trends of change in GDI, KBM, and CBI. For remote sites, the proportions of the three HCI methods are basically similar, specifically, with KBM accounting for the highest percentage at 36.6%, followed closely by GDI at 33.8% and CBI at 29.6% (see Fig. 14(a)), respectively. For local sites, compared with the interaction methods on the local side, it is evident that GDI dominates in all three categories annually(see Fig. 13), with respective proportions of 75% for GDI, 13.2% for KBM, and 11.8% for CBI(see Fig. 14(b)). This suggests that in remote collaboration scenarios, there isn't a clear dominance of any single interaction method. Instead, a variety of methods are utilized, possibly reflecting the diverse needs and preferences of remote users. The similarity in proportions might also indicate that each method has its own strengths and weaknesses, and thus, a combination of methods is employed to optimize the remote collaboration experience. The analysis of interaction methods in XR remote collaboration from the perspective of supporting visual annotations cues in industry reveals distinct patterns at remote and local sites. While remote sites exhibit a more balanced utilization of GDI, KBM, and CBI, local sites are dominated by GDI. The stability in the proportions of these methods suggests a need for further From Table 2, we summarize the essential toolkits used in XR remote collaboration for industry, focusing on those that support visual annotation cues. These toolkits enable researchers and industry professionals to develop prototype systems for practical applications like industrial assembly, training, and maintenance. investigation into the factors influencing interaction method preferences and the potential benefits of exploring new interaction methods to improve XR remote collaboration.



(a)HCI in remote sites



(b)HCI in local sites

Fig. 13: The trend of HCI in XR remote collaboration.

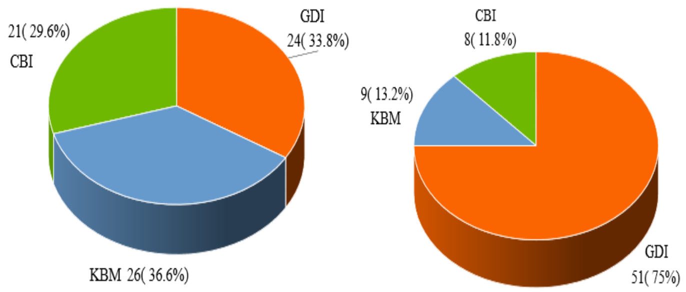


Fig. 14: The proportion of interaction modalities.

4.8 XR toolkits

The primary VR toolkits that are widely utilized in the domain of XR visualization and remote collaboration include VRTK (Virtual Reality Toolkit), which offers a comprehensive set of pre-built components and interactions to streamline the development process; SteamVR, providing robust support for a variety of VR hardware and enabling developers to create immersive experiences with relative ease; the Oculus XR Interaction Toolkit, designed specifically for Oculus devices and offering optimized interactions and gestures that enhance user engagement within virtual environments; and the Meta Avatars SDK, which allows for the creation of personalized and expressive avatars, adding a human touch to VR interactions and making virtual collaboration more intuitive and engaging.

For AR/MR, several key toolkits stand out as vital resources for developers and researchers aiming to create immersive and interactive applications. Vuforia is renowned for its sophisticated image target recognition and tracking capabilities, allowing for the seamless overlay of digital information onto the physical world. ARToolkit, as a pioneer in the field, provides a robust foundation for building augmented reality experiences with its efficient marker-based tracking system. Vuforia Chalk takes collaboration to new heights by enabling real-time, interactive annotations in augmented reality environments, facilitating effective communication and problem-solving among remote teams. Webkit, with its powerful rendering engine, supports the development of web-based AR applications that can be accessed across various devices, promoting wider accessibility. Apple's ARKit and Google's ARCore are platform-specific solutions that leverage the capabilities of IOS and Android devices, respectively, offering advanced features like environmental understanding and motion tracking to create realistic AR experiences. AR tags serve as simple yet effective markers that can be easily recognized by AR systems, providing a stable basis for overlaying digital content. Lastly, MRTK is a comprehensive suite designed specifically for Microsoft's HoloLens and other XR platforms, offering a wide array of tools and components to simplify the development process and enable the creation of rich, interactive XR applications.

In the field of telecommunication for XR remote collaboration, a variety of toolkits play crucial roles. More specifically, WebRTC, supported by major browser vendors, enables peer-to-peer real-time communication, including video, voice, and data transfer, making it ideal for creating immersive communication solutions. The WebSocket API allows for persistent, two-way interactive communication between a user's browser and a server, enabling real-time updates and interactions without the need for continuous polling. Photon Unity Networking is particularly popular in game development for its efficient handling of multiplayer interactions, which can be adapted for collaborative XR environments. Agora's real-time network provides scalable and low-latency communication solutions, essential for smooth and interactive remote collaboration experiences. Socket.io library offers flexible real-time communication capabilities, supporting both WebSockets and HTTP long polling to ensure compatibility across different network conditions. Additionally, Dynamics 365 Remote Assist and Teams integrate telecommunication features with remote assistance functionalities, enabling hands-free communication and collaboration in industrial settings. Zoom and Microsoft Teams are widely used video conferencing tools that can also be leveraged for XR remote collaboration scenarios. Furthermore, VRChat SDK allows developers to create social and collaborative VR experiences, emphasizing user interaction and presence. Unity UNet API provides networking capabilities for Unity-based applications, facilitating synchronization and communication between multiple users in XR environments. These toolkits collectively provide the necessary infrastructure and functionalities to support real-time, interactive, and immersive telecommunication in XR applications, enabling seamless remote collaboration across various industries.

In addition, we compared commercial XR platforms (*e.g.*, Microsoft HoloLens, Magic Leap, Meta Quest) in terms of technical performance and user experience as shown in Table 4.

Finally, deep learning algorithms include YOLOv5, Segment Anything, and Instant-NGP. The main VR devices are HTC Vive and Oculus, while the AR/MR devices are HoloLens 1/2.

5. Discussion

After carefully reviewing 69 publications within this domain (referenced in Table 2), it has been observed that industrial applications predominantly revolve around key areas such as design, assembly, inspection, training, maintenance, and beyond. In this section, a thorough exploration of research pertaining to XR remote collaboration, with an emphasis on the utilization of visual annotation cues within industrial contexts, will also be comprehensively discussed.

5.1 Collected studies

While our review primarily centers on XR remote collaboration utilizing visual annotation cues within industrial contexts, we have also gathered several works that explore XR remote collaborative applications in healthcare settings (*e.g.*, [26,39,41,48,50,58,74]) supporting annotations. Additionally, we included "ShareYourReality",[38] a notable example of VR remote collaboration, which has no ability to share annotations. These works have been incorporated into our review because they can provide valuable insights into supporting visual annotation cues and highlight the key technologies that merit consideration in such systems. Although we have carried out an extensive survey, it is possible that some relevant research may not have been included in our review due to time constraints.

5.2 Document type and conducting studies

Our analysis reveals a predominant preference among researchers for journal-based publications, with the IJAMT emerging as the most frequent outlet (see Fig. 5). However, our investigation identified limited relevant publications (n=3 from ISMAR conference proceedings) on XR remote collaboration utilizing visual annotations, primarily appearing in specialized conferences including IEEE VR, VRST, UIST, and CSCW. Notably, the CHI conference demonstrates the highest publication count in this domain (see Fig. 6). This scarcity of related studies aligns with previous systematic reviews[2,96] documenting the under-explored nature of collaborative XR implementations. Thematic analysis indicates that HCI constitutes the predominant research focus, representing over four-fifths (81%) of examined publications. For comprehensive guidance on HCI research methodologies, we recommend consulting the seminal work by Lazar *et al*,[97] which provides extensive coverage of established techniques, critical evaluations, and practical implementation strategies for HCI investigations.

5.3 User interfaces and HCI

The design challenge for user interfaces is optimizing HCI usability without overcrowding while preserving natural interaction.[98] Based on Fig. 13, and Fig. 14, the lack of obvious trends of change in the proportions of GDI, KBM, and CBI over time, which indicates a certain stability in the interaction methods employed in XR remote collaboration for remote sites. This stability might be due to the maturity of these methods or the absence of significant technological advancements that could alter the landscape of interaction methods in this domain. However, it is also possible that the stability reflects a lack of exploration or adoption of new interaction methods, which could potentially enhance the effectiveness of XR remote collaboration. As shown in Fig. 9(a) and Fig. 10(a), remote sites rely primarily on VR HMDs (49%) and desktop systems (37%) for XR visualization. This aligns with distinct HCI patterns: VR collaborative environments using HMDs favor GDI and CBI, while desktop displays predominantly employ KBM. Statistical analysis reveals a strong correlation between display modes and interaction methods.

Table 4: Commercial XR platforms in terms of technical performance and user experience.

	Dimension	HoloLens 2	Magic Leap 2	Meta Quest 3
Technical performance Dimension	Field of View (FOV)	~52° (limited by waveguide optics)	~70° (dynamic focus)	~110° (Pancake lenses)
	Display Technology	Waveguide (3-layer diffractive)	Light-field (6-layer waveguide)	LCD (120 Hz refresh)
	Tracking Accuracy	High (depth sensors + IMU)	Moderate (environmental drift issues)	High (inside-out + AI)
	Compute Platform	Snapdragon 850 (on-device)	NVIDIA Tegra X2 (lightpack)	Snapdragon XR2 Gen 2
	Latency	15–20 ms (render-to-photon)	<30 ms (cloud-assisted)	<20 ms (standalone)
User Experience Dimension	Comfort	Moderate (front-heavy, 645g)	High (split design, 260g HMD)	Moderate (503g, strap-dependent)
	Interaction	Hand gestures (high accuracy)	Controller + gestures (inconsistent)	Hands/controllers (flexible but tracking drops)
	Content Ecosystem	Enterprise-focused (UWP)	Limited (LuminOS)	Rich (App Lab + SteamVR)
	Visual Comfort	Moderate (fixed focal plane)	High (dynamic focus)	Moderate (fixed)

In contrast, for local sites, the stark difference in proportions compared to remote sites suggests that local users may have different interaction preferences or that the local environment may be more conducive to certain types of interaction methods. The dominance of GDI at local sites could be attributed to factors such as familiarity, ease of use, or the specific requirements of the local tasks. Importantly, GDI better facilitates user-created visual annotation cues, enhancing collaborative clarity. The prevalence of GDI at local sites may stem from user familiarity, task-specific demands, or ergonomic advantages. Notably, GDI enhances collaborative clarity by enabling intuitive user-generated visual annotations. As evidenced in Fig. 9(b) and Fig. 10(b), local sites favor MR HMDs (40%) and HHDs (27%) for on-site XR workflows. This aligns with gesture interaction being the optimal input method for MR and handheld devices, a correlation strongly supported by the statistical data.

5.4 Creating accurate annotations

XR remote collaboration using visual annotations in industry faces a key challenge: creating accurate annotations. GDI methods, while freehand gestures enable natural annotation creation, are hindered by algorithm limitations, sensor issues, and noise, causing misinterpretation of fine movements and poor sketch quality. Additionally, CBI-based annotations, though precise, restrict natural creativity and pose ergonomic barriers. These technical gaps frequently yield misaligned or ambiguous annotations, escalating error risks in precision-dependent domains like manufacturing, prototyping, and maintenance. Mitigation demands hybrid interfaces combining adaptive gesture recognition (enhanced via AI-driven motion prediction) with context-aware controller assistance, coupled with multi-sensory feedback systems to bridge the virtual-physical annotation divide.

5.5 Core theoretical frameworks

Based on the systematic scoping survey, in industrial XR remote collaboration, situational awareness and embodied interaction form a synergistic relationship, amplified by visual annotation cues, to enhance distributed workflow efficiency. Situational awareness refers to users' perception, comprehension, and projection of environmental states (*e.g.*, machinery status, collaborator actions), while embodied interaction emphasizes physical movement and gestures as primary interaction modes. Visual annotations, such as AR overlays, 3D markers, or VR spatial guides, serve as the critical bridge between these two concepts.

The interplay is bidirectional: embodied interaction enhances situational awareness by linking natural actions (*e.g.*, gestures, gaze) to real-time contextual data, reducing cognitive load. For instance, a technician's hand motion to rotate a virtual engine part triggers AR annotations displaying maintenance steps, improving spatial comprehension. Conversely, situational awareness drives embodied interaction design by dynamically adjusting interfaces based on contextual needs. For example, if situational awareness detects a misaligned robotic arm, the system may pre-load virtual adjustment handles to guide embodied corrections.

5.6 Virtual-real fusion of annotations

We find that both spatial anchoring and computer-vision-based registration are used to guarantee that annotations remain correctly aligned between the remote expert's and the local worker's views. Specifically, spatial anchoring (world-stabilized annotations) is the dominant mechanism. Most systems place annotations relative to 3-D feature points or physical markers that are tracked by the on-site device's SLAM pipeline. This keeps an arrow, highlight, or free-hand stroke "stuck" to the same bolt, cable, or panel even when the worker moves around the object.

Additionally, computer-vision-based registration is employed as an additional safeguard or fallback. When the on-site camera view changes abruptly (*e.g.*, the user steps back), simple spatial anchoring can drift. Several studies therefore combine SLAM with lightweight CV techniques—such as keypoint re-detection or planar homography refinement—to re-lock the annotation to the intended surface. Marques *et al.*,[78] for example, allow the remote expert to manually freeze the incoming video, draw an annotation on the static frame, and then release the frame; this hybrid CV-plus-anchor method prevents mis-registration when the local user's viewpoint shifts while the expert is drawing.

In short, spatial anchoring provides the primary synchronization layer, while computer-vision-based registration is used to correct or re-anchor annotations in case of tracking failure or viewpoint drift.

6. Future directions

6.1 XR dynamic visualization

In XR remote collaboration, scene perception has conventionally been based on video and audio.[2,8,10,21] Nevertheless, we summarized that researchers are currently incorporating 3D reconstruction to supply depth information and more comprehensive environmental details, thereby improving the comprehension of remote surroundings.[13,15,20–22] From a visualization standpoint, the difficulty lies in optimizing visual realism while guaranteeing correct understanding and promoting trust, all the while curtailing the time users spend in XR and circumventing misplaced confidence. A pivotal element in reconciling these trade-offs is the level of detail (LOFD) utilized to present objects within the 3D XR scene. However, in the current study, researchers basically did not consider one point, and the focus was on how to create visual annotation cues efficiently. Dynamic LOFD adjusts rendering levels: high detail for critical XR elements near users, lower for distant/less relevant ones, optimizing computation and enhancing collaborative focus. This method facilitates more effective transmission of scene and task-related information, enabling participants to concentrate on critical elements without being inundated by superfluous details.

6.2 HCI using intelligent agent

Agent-driven architectures in HCI are advancing at an accelerated pace, spanning varied implementation domains and scholarly investigations. These intelligent platforms prioritize delivering streamlined and user-centric interaction frameworks while tackling critical issues, including privacy preservation and contextual adaptability.[99] The integration of intelligent agents into HCI for industrial XR remote collaboration presents transformative potential to address persistent challenges in visual annotations. Building on the foundation of agent-driven architectures, future research should prioritize context-aware adaptive interfaces that dynamically optimize annotation. Intelligent agents could leverage real-time user intent analysis, via multimodal inputs (gestures, gaze, voice), to auto-calibrate gesture recognition thresholds in XR environments, mitigating current limitations in fine motor skill capture during technical sketching. Recently, Chen *et al*.[100] reveals how large language model integration revolutionizes agent-based systems, creating unprecedented capabilities in autonomous learning paradigms and decision-making processes. Thus, intelligent agent-driven HCI could interpret ambiguous visual cues through semantic grounding, automatically generating context-appropriate annotations while preserving collaborators' intent across distributed teams, reducing cognitive load while maintaining natural interaction. Ultimately, intelligent agent-driven HCI must evolve as contextual mediators, balancing industrial tasks' requirements with ergonomic interaction paradigms to unlock XR's full potential in next-generation industrial remote collaboration.

6.3 Embodied intelligent robot-driven remote collaboration

Object-centric XR remote collaboration is increasingly critical in industrial maintenance and assembly, where distributed teams co-analyze and troubleshoot physical tasks using visual annotation cues. Despite its growing adoption for enabling collaboration, current systems lack flexible ways of supporting remote collaborative tasks in industrial settings. While contemporary XR collaborative platforms partially address these needs through shared telepresence, they suffer from two critical industrial limitations: (1) rigid viewpoint constraints that prevent context-aware annotation positioning around multi-scale equipment, (2) inability to synchronize perspective-dependent annotations across varying HMD/HHD displays.

To overcome these challenges, embodied intelligent robots and computer vision systems could revolutionize industrial XR remote collaboration by serving as dynamic physical proxies at local worksites. By deploying embodied intelligent robots, remote collaborators can gain unprecedented flexibility to navigate complex environments, dynamically adjust viewpoints around multi-scale equipment, and contextually anchor visual annotations to physical objects. This breaks the rigid scene perception perspective constraints while enabling perspective-aware annotation synchronization across displays through real-time spatial mapping. As Feick *et al*.[101] suggested, combining robotic mobility allows the system to share the right scene perspective. As AI advances, these embodied agents will evolve from passive teleoperation tools to proactive collaborators that predict needs through behavioral understanding, fundamentally redefining human-machine partnership in Industry 5.0 environments.

7. Conclusion and perspectives

This systematic review investigates the implementation and impact of visual annotation systems in industrial XR remote collaboration, synthesizing 69 studies (2019-2024). To our knowledge, it's the first to offer a clear, concise map of XR-based remote collaboration with visual annotations in industry. XR remote collaboration has seen significant industrial progress, with broad design, assembly, and maintenance applications, and sustained academic interest. We aim to provide a foundation and guide for researchers interested in this field.

Moving XR annotation research from labs to engineering and widespread use presents opportunities and challenges. Fortunately, as more researchers focus on this area, engaging in industrial XR remote collaboration with visual annotations is becoming increasingly accessible. Looking ahead, there are several promising directions for future research such as XR dynamic visualization, HCI using intelligent agent, and embodied intelligent robot-driven remote collaboration. Future studies should focus on addressing the scalability and cost-effectiveness issues to facilitate the widespread implementation of XR remote collaboration with visual annotations in various industries.

Acknowledgments

This research was financially sponsored by the National Natural Science Foundation of China (No. 52405533), and National Key Research and Development Program of China(2022YFB3306100). Besides, we would like to thank the anonymous reviewers for their constructive suggestions for enhancing this paper.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

CRediT Statement

Peng Wang: Writing - Original draft, Conceptualization, Resources, Formal analysis, **Yue Wang:** Formal analysis, Investigation, **Yiwei Wang:** Formal analysis, Investigation, Supervision, **Mark Billinghurst:** Supervision and Methodology, **Huan Yang:** Data curation, Visualization, Formal analysis, Investigation, **Dongyu Yang:** Resources, Data curation, Formal analysis, Investigation, **Rong Luo** and **Xiangyu Zhang:** Formal analysis, Investigation.

References

[1] A. Schäfer, G. Reis, D. Stricker, A survey on synchronous augmented, virtual, and Mixed reality remote collaboration systems, *ACM Computing Surveys*, 2023, **55**, 1-27, doi: 10.1145/3533376.

[2] P. Wang, X. Bai, M. Billinghurst, S. Zhang, X. Zhang, S. Wang, W. He, Y. an, H. Ji, AR/MR remote collaboration on physical tasks: a review, *Robotics and Computer-Integrated Manufacturing*, 2021, **72**, 102071, doi: 10.1016/j.rcim.2020.102071.

[3] S. Feng, W. He, X. Zhang, M. Billinghurst, S. Wang, A comprehensive survey on AR-enabled local collaboration, *Virtual Reality*, 2023, **27**, 2941-2966, doi: 10.1007/s10055-023-00848-2.

[4] S. Zhang, Y. Li, K. L. Man, J. S. Smith, Y. Yue, Understanding user experience, task performance, and task interdependence in symmetric and asymmetric VR collaborations, *Virtual Reality*, 2024, **29**, 6, doi: 10.1007/s10055-024-01072-2.

[5] J. Pang, P. Zheng, J. Fan, T. Liu, Towards cognition-augmented human-centric assembly: a visual computation perspective, *Robotics and Computer-Integrated Manufacturing*, 2025, **91**, 102852, doi: 10.1016/j.rcim.2024.102852.

[6] Z. Wang, X. Bai, S. Zhang, M. Billinghurst, W. He, P. Wang, W. Lan, H. Min, Y. Chen, A comprehensive review of augmented reality-based instruction in manual assembly, training and repair, *Robotics and Computer-Integrated Manufacturing*, 2022, **78**, 102407, doi: 10.1016/j.rcim.2022.102407.

[7] M. Eswaran, A. K. Gulivindala, A. K. Inkulu, M. V. A. Raju Bahubalendruni, Augmented reality-based guidance in product assembly and maintenance/repair perspective: a state of the art review on challenges and opportunities, *Expert Systems with Applications*, 2023, **213**, 118983, doi: 10.1016/j.eswa.2022.118983.

[8] P. Wang, X. Bai, M. Billinghurst, S. Zhang, S. Wei, G. Xu, W. He, X. Zhang, J. Zhang, 3DGAM: using 3D gesture and CAD models for training on mixedreality remote collaboration, *Multimedia Tools and Applications*, 2021, **80**, 31059-31084, doi: 10.1007/s11042-020-09731-7.

[9] Z. Wang, L. Li, Y. Liu, Y. Jiang, Y. Wang, Y. Dai, An experimental study on mixed reality-based user interface for collaborative operation of high-precision process equipment, *The International Journal of Advanced Manufacturing Technology*, 2024, **132**, 2443-2459, doi: 10.1007/s00170-024-13517-8.

[10] X. Zhang, X. Bai, S. Zhang, W. He, S. Wang, Y. Yan, P. Wang, L. Liu, A novel mixed reality remote collaboration system with adaptive generation of instructions, *Computers & Industrial Engineering*, 2024, **194**, 110353, doi: 10.1016/j.cie.2024.110353.

[11] Y. Yan, X. Bai, W. He, S. Wang, X. Zhang, L. Liu, Q. Yu, B. Zhang, Less gets more attention: a novel human-centered MR remote collaboration assembly method with information recommendation and visual enhancement, *Robotics and Computer-Integrated Manufacturing*, 2025, **92**, 102898, doi: 10.1016/j.rcim.2024.102898.

[12] P. Wang, S. Zhang, M. Billinghurst, X. Bai, W. He, S. Wang, M. Sun, X. Zhang, A comprehensive survey of AR/MR-based co-design in manufacturing, *Engineering with Computers*, 2020, **36**, 1715-1738, doi: 10.1007/s00366-019-00792-3.

[13] C. G. Fidalgo, Y. Yan, H. Cho, M. Sousa, D. Lindlbauer, J. Jorge, A survey on remote assistance and training in mixed reality environments, *IEEE Transactions on Visualization and Computer Graphics*, 2023, **29**, 2291-2303, doi: 10.1109/TVCG.2023.3247081.

[14] A. de Giorgio, F. M. Monetti, A. Maffei, M. Romero, L. Wang, Adopting extended reality? A systematic review of manufacturing training and teaching applications, *Journal of Manufacturing Systems*, 2023, **71**, 645-663, doi: 10.1016/j.jmsy.2023.10.016.

[15] B. Wang, L. Zheng, Y. Wang, W. Fang, L. Wang, Towards the industry 5.0 frontier: review and prospect of XR in product assembly, *Journal of Manufacturing Systems*, 2024, **74**, 777-811, doi: 10.1016/j.jmsy.2024.05.002.

[16] X. Wang, L. Shen, L.-H. Lee, A systematic review of XR-enabled remote human-robot interaction systems, *ACM Computing Surveys*, 2025, **57**, 1-37, doi: 10.1145/3730574.

[17] M. Mateescu, C. Pimmer, C. Zahn, D. Klinkhammer, H. Reiterer, Collaboration on large interactive displays: a systematic review, *Human–Computer Interaction*, 2021, **36**, 243-277, doi: 10.1080/07370024.2019.1697697.

[18] W. Fang, L. Chen, T. Zhang, C. Chen, Z. Teng, L. Wang, Head-mounted display augmented reality in manufacturing: a systematic review, *Robotics and Computer-Integrated Manufacturing*, 2023, **83**, 102567, doi: 10.1016/j.rcim.2023.102567.

[19] M. Eswaran, M. V. A. R. Bahubalendruni, Challenges and opportunities on AR/VR technologies for manufacturing systems in the context of industry 4.0: a state of the art review, *Journal of Manufacturing Systems*, 2022, **65**, 260-278, doi: 10.1016/j.jmsy.2022.09.016.

[20] L. Gao, Using mixed reality for asymmetric remote collaboration in a room-scale workspace, *Human Interface Technology,* 2020, doi: 10.26021/10814.

[21] B. Marques, Concepts and methods to support the development and evaluation of remote collaboration using augmented reality, Doctoral Dissertation,University of Aveiro, 2021, doi: 10.2312/2633268.

[22] T. Teo, Techniques for the Immersive Remote Collaboration Combining 360° Video with 3D Reconstruction, Doctoral Dissertation, University of South Australia*,* 2020.

[23] Z. Wang, X. Zhang, L. Li, Y. Zhou, Z. Lu, Y. Dai, C. Liu, Z. Su, X. Bai, M. Billinghurst, Evaluating visual encoding quality of a mixed reality user interface for human–machine co-assembly in complex operational terrain, *Advanced Engineering Informatics*, 2023, **58**, 102171, doi: 10.1016/j.aei.2023.102171.

[24] X. Zhang, X. Bai, S. Zhang, W. He, S. Wang, Y. Yan, Q. Yu, L. Liu, A novel MR remote collaboration system using 3D spatial area cue and visual notification, *Journal of Manufacturing Systems*, 2023, **67**, 389-409, doi: 10.1016/j.jmsy.2023.02.013.

[25] J. Shin, G. Suh, S.-H. Kim, H.-J. Yang, J.-E. Shin, G. Lee, S.

Kim, Effect of the hand and gaze pointers in remote collaboration, *IEEE Access*, 2024, **12**, 172774-172784.

[26] V. Chheang, D. Schott, P. Saalfeld, L. Vradelis, T. Huber, F. Huettl, H. Lang, B. Preim, C. Hansen, Advanced liver surgery training in collaborative VR environments, *Computers & Graphics*, 2024, **119**, 103879, doi: 10.1016/j.cag.2024.01.006.

[27] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement, *International Journal of Surgery*, 2010, **8**, 336-341, doi: 10.1016/j.ijsu.2010.02.007.

[28] M. L. Rethlefsen, M. J. Page, PRISMA 2020 and PRISMA-S: common questions on tracking records and the flow diagram, *Journal of the Medical Library Association*, 2022, **110**, 253-257, doi: 10.5195/jmla.2022.1449.

[29] MJ, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, *et al*, The PRISMA 2020 statement: An updated guideline for reporting systematic reviews, *International Journal of Surgery*, 2021, **372**, n71, doi: 10.1136/bmj.n71.

[30] A. Dey, M. Billinghurst, R. W. Lindeman, J. E. Swan II, A systematic review of usability studies in augmented reality between 2005 and 2014, *IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, Merida, Mexico, September 19-23, 2016, 49-50, doi: 10.1109/ISMAR-Adjunct.2016.0036.

[31] M. Gattullo, A. Evangelista, A. E. Uva, M. Fiorentino, J. L. Gabbard, What, how, and why are visual assets used in industrial augmented reality? a systematic review and classification in maintenance, assembly, and training from 1997 to 2019, *IEEE Transactions on Visualization and Computer Graphics*, 2022, **28**, 1443-1456, doi: 10.1109/TVCG.2020.3014614.

[32] E. Chang, Y. Lee, M. Billinghurst, B. Yoo, Efficient VR-AR communication method using virtual replicas in XR remote collaboration, *International Journal of Human-Computer Studies*, 2024, **190**, 103304, doi: 10.1016/j.ijhcs.2024.103304.

[33] N. Nugegoda, M.-E. Jannat, K. Hasan, P. Lasserre, Exploring the effect of viewing attributes of mobile AR interfaces on remote collaborative and competitive tasks, *IEEE Transactions on Visualization and Computer Graphics*, 2024, **30**, 7288-7298, doi: 10.1109/TVCG.2024.3456177.

[34] B. Marques, C. Ferreira, S. Silva, A. Santos, A. Santos, P. Dias, B. S. Santos, Exploring different content creation and display methods for remote collaboration supported by eXtended reality: comparative analysis of distinct task scenarios, *Multimedia Tools and Applications*, 2025, **84**, 19785-19815, doi: 10.1007/s11042-024-19836-y.

[35] J. E. S. Grønbæk, J. Sánchez Esquivel, G. Leiva, E. Velloso, H. Gellersen, K. Pfeuffer, Blended whiteboard: physicality and reconfigurability in remote mixed reality collaboration, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Honolulu, HI, USA, 2024, 1-16., doi: 10.1145/3613904.3642293.

[36] E. Wong, J. E. S. Grønbæk, E. Velloso, The jamais vu effect: understanding the fragile illusion of co-presence in mixed reality, *Designing Interactive Systems Conference,* Copenhagen, Denmark, 2024, 2227-2246, doi: 10.1145/3643834.3661574.

[37] E. Wong, J. Sánchez Esquivel, G. Leiva, J. E. S. Grønbæk, E. Velloso, Practice-informed patterns for organising large groups in distributed mixed reality collaboration, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Honolulu, HI, USA, 2024, 1-18, doi: 10.1145/3613904.3642502.

[38] K. P. Venkatraj, W. Meijer, M. Perusquia-Hernandez, G. Huisman, A. El Ali, ShareYourReality: investigating haptic feedback and agency in virtual avatar co-embodiment, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2024, 1-15, doi: 10.1145/3613904.3642425.

[39] S. Lambert, S. Voros, G. Canlorbe, J. Troccaz, I. Avellino, Understanding takeovers and telestration in laparoscopic surgery to inform telementoring system design, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Honolulu, HI, USA, 2024, 1-17, doi: 10.1145/3613904.3641978.

[40] T. Rasmussen, K. Grønbæk, W. Huang, Supporting tailorability in augmented reality based remote assistance in the manufacturing industry: a user study, *Computers & Graphics*, 2024, **125**, 104095, doi: 10.1016/j.cag.2024.104095.

[41] Z. Lu, J. H. Seo, P. Khandan, T. Maxa, B. Garcia-Pi, E. Wells-Beede, The work behind an interdisciplinary team: creating a postpartum hemorrhage virtual reality training platform, *Clinical Simulation in Nursing*, 2024, **94**, 101595, doi: 10.1016/j.ecns.2024.101595.

[42] X. Huang, M. Yin, Z. Xia, R. Xiao, VirtualNexus: enhancing 360-degree video AR/VR collaboration with environment cutouts and virtual replicas, *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology,* Pittsburgh, USA, 2024, 1-12, doi: 10.1145/3654777.3676377.

[43] N.-J. Shih, Y.-C. Wu, Augmented reality- and geographic information system-based inspection of brick details in heritage warehouses, *Applied Sciences*, 2024, **14**, 8316, doi: 10.3390/app14188316.

[44] V. O'Keeffe, R. Jang, K. Manning, R. Trott, S. Howard, A.-L. Hordacre, J. Spoehr, Forming a view: a human factors case study of augmented reality collaboration in assembly, *Ergonomics*, 2024, **67**, 1828-1844, doi: 10.1080/00140139.2024.2352733.

[45] K. Seki, S. Fukushige, TeleGhost: asymmetric telepresence system using AR and VR avatars in a shared real space, *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Honolulu, HI, USA, October 1-4, 2023, 4550-4555, doi: 10.1109/SMC53992.2023.10394515.

[46] X. Zhang, X. Bai, S. Zhang, W. He, P. Wang, Z. Wang, Y. Yan, Q. Yu, Real-time 3D video-based MR remote collaboration using gesture cues and virtual replicas, *The International Journal of Advanced Manufacturing Technology*, 2022, **121**, 7697-7719, doi: 10.1007/s00170-022-09654-7.

[47] Y. Yan, X. Bai, W. He, S. Wang, X. Zhang, P. Wang, L. Liu, B. Zhang, Can you notice my attention? A novel information vision enhancement method in MR remote collaborative assembly, *The International Journal of Advanced Manufacturing Technology*, 2023, **127**, 1835-1857, doi: 10.1007/s00170-023-11652-2.

[48] J. Liu, K. Qian, Z. Qin, M. D. Alshehri, Q. Li, Y. Tai, Cloud computing-enabled IIOT system for neurosurgical simulation using augmented reality data access, *Digital Communications and Networks*, 2023, **9**, 347-357, doi: 10.1016/j.dcan.2022.04.019.

[49] J. Seo, I. Avellino, D. P. Rajasagi, A. Komlodi, H. M. Mentis, HoloMentor, *Proceedings of the ACM on Human-Computer Interaction*, 2023, **7**, 1-29, doi: 10.1145/3567561.

[50] K. Yu, D. Roth, R. Strak, F. Pankratz, J. Reichling, C. Kraetsch, S. Weidert, M. Lazarovici, N. Navab, U. Eck, Mixed reality 3D teleconsultation for emergency decompressive craniotomy: an evaluation with medical residents, *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Sydney, Australia, October 16-20, 2023, 662-671, doi: 10.1109/ISMAR59233.2023.00081.

[51] P. Praveena, Y. Wang, E. Senft, M. Gleicher, B. Mutlu, Periscope: a robotic camera system to support remote physical collaboration, *Proceedings of the ACM on Human-Computer Interaction*, 2023, **7**, 1-39, doi: 10.1145/3610199.

[52] H. Meng, Y. Wang, P. Praveena, M. Gleicher, B. Mutlu, Demonstrating periscope: a robotic camera system to support remote physical collaboration, *Computer Supported Cooperative Work and Social Computing,* Minneapolis, USA, 2023, 365-368, doi: 10.1145/3584931.3607493.

[53] J. E. S. Grønbæk, K. Pfeuffer, E. Velloso, M. Astrup, M. I. S. Pedersen, M. Kjær, G. Leiva, H. Gellersen, Partially blended realities: aligning dissimilar spaces for distributed mixed reality meetings, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Hamburg, Germany, 2023, 1-16, doi: 10.1145/3544548.3581515.

[54] H. Tian, G. A. Lee, H. Bai, M. Billinghurst, Using virtual replicas to improve mixed reality remote collaboration, *IEEE Transactions on Visualization and Computer Graphics*, 2023, **29**, 2785-2795, doi: 10.1109/TVCG.2023.3247113.

[55] M. Augstein, T. Neumayr, J. Schönböck, C. Kovacs, Remote persons are closer than they appear: home, team and a lockdown, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Hamburg, Germany, 2023, 1-25, doi: 10.1145/3544548.3580989.

[56] S. Kim, W. Huang, C.-M. Oh, G. Lee, M. Billinghurst, S.-J. Lee, View types and visual communication cues for remote collaboration, *Computers, Materials & Continua*, 2023, **74**, 4363-4379, doi: 10.32604/cmc.2023.034209.

[57] F. Zaman, C. Anslow, A. Chalmers, T. Rhee, MRMAC: mixed reality multi-user asymmetric collaboration, *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Sydney, Australia, October 16-20, 2023, 591-600, doi: 10.1109/ISMAR59233.2023.00074.

[58] S. Maria, H. M. Mentis, G. Canlorbe, I. Avellino, Supporting collaborative discussions in surgical teleconsulting through augmented reality head mounted displays, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Hamburg, Germany, 2023, 1-13., doi: 10.1145/3544548.3580714.

[59] Z. Wang, Y. Wu, V. A. González, Y. Zou, E. del Rey Castillo, M. Arashpour, G. Cabrera-Guerrero, User-centric immersive virtual reality development framework for data visualization and decision-making in infrastructure remote inspections, *Advanced Engineering Informatics*, 2023, **57**, 102078, doi: 10.1016/j.aei.2023.102078.

[60] B. Marques, C. Ferreira, S. Silva, P. Dias, B. S. Santos, Is social presence (alone) a general predictor for good remote collaboration? comparing video and augmented reality guidance in maintenance procedures, *Virtual Reality*, 2023, **27**, 1783-1796, doi: 10.1007/s10055-023-00770-7.

[61] J. Wang, X. Fan, Y. Zhu, X. Yang, X. Yin, Cross-platform AR annotation for assembly-design communication in pipe outfitting, *The International Journal of Advanced Manufacturing Technology*, 2022, **121**, 4981-4998, doi: 10.1007/s00170-022-09637-8.

[62] H. M. Mentis, I. Avellino, J. Seo, AR HMD for remote instruction in healthcare, *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Christchurch, New Zealand, March 12-16, 2022, 437-440, doi: 10.1109/VRW55335.2022.00096.

[63] A. Jing, K. Gupta, J. McDade, G. A. Lee, M. Billinghurst, Comparing gaze-supported modalities with empathic mixed reality interfaces in remote collaboration, *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Singapore, October 17-21, 2022, 837-846, doi: 10.1109/ISMAR55827.2022.00102.

[64] C.-H. Chae, N.-J. Jung, K.-H. Ko, Mobile power facilities maintenance system using augmented reality, *Journal of Electrical Engineering & Technology*, 2022, **17**, 1357-1369, doi: 10.1007/s42835-021-00942-y.

[65] Y. Wang, P. Wang, Z. Luo, Y. Yan, A novel AR remote collaborative platform for sharing 2.5D gestures and gaze, *The International Journal of Advanced Manufacturing Technology*, 2022, **119**, 6413-6421, doi: 10.1007/s00170-022-08747-7.

[66] S. Alirezaei, H. Taghaddos, K. Ghorab, A. N. Tak, S. Alirezaei, BIM-augmented reality integrated approach to risk management, *Automation in Construction*, 2022, **141**, 104458, doi: 10.1016/j.autcon.2022.104458.

[67] B. Marques, S. Silva, A. Teixeira, P. Dias, B. S. Santos, A vision for contextualized evaluation of remote collaboration supported by AR, *Computers & Graphics*, 2022, **102**, 413-425, doi: 10.1016/j.cag.2021.10.009.

[68] B. Marques, S. Silva, J. Alves, A. Rocha, P. Dias, B. S. Santos, Remote collaboration in maintenance contexts using augmented reality: insights from a participatory process, *International Journal on Interactive Design and Manufacturing (IJIDeM*), 2022, **16**, 419-438, doi: 10.1007/s12008-021-00798-6.

[69] T. Madeira, B. Marques, J. Alves, P. Dias, B. S. Santos, Exploring annotations and hand tracking in augmented reality for remote collaboration, *Human Systems Engineering and Design III*, Cham: Springer International Publishing, 2020, 83-89, doi: 10.1007/978-3-030-58282-1_14.

[70] D. Calandra, A. Cannavò, F. Lamberti, Improving AR-powered remote assistance: a new approach aimed to foster operator's autonomy and optimize the use of skilled resources, *The International Journal of Advanced Manufacturing*

Technology, 2021, **114**, 3147-3164, doi: 10.1007/s00170-021-06871-4.

[71] Y. Lee, B. Yoo, XR collaboration beyond virtual reality: work in the real world, *Journal of Computational Design and Engineering*, 2021, **8**, 756-772, doi: 10.1093/jcde/qwab012.

[72] T. Ludwig, O. Stickel, P. Tolmie, M. Sellmer, shARe-IT: ad hoc remote troubleshooting through augmented reality, *Computer Supported Cooperative Work (CSCW)*, 2021, **30**, 119-167, doi: 10.1007/s10606-021-09393-5.

[73] S. Tea, K. Panuwatwanich, R. Ruthankoon, M. Kaewmoracharoen, Multiuser immersive virtual reality application for real-time remote collaboration to enhance design review process in the social distancing era, *Journal of Engineering, Design and Technology*, 2022, **20**, 281-298, doi: 10.1108/jedt-12-2020-0500.

[74] D. Gasques, J. G. Johnson, T. Sharkey, Y. Feng, R. Wang, Z. R. Xu, E. Zavala, Y. Zhang, W. Xie, X. Zhang, K. Davis, M. Yip, N. Weibel, ARTEMIS: a collaborative mixed-reality system for immersive surgical telementoring, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan, 2021, 1-14, doi: 10.1145/3411764.3445576.

[75] K. Yu, G. Gorbachev, U. Eck, F. Pankratz, N. Navab, D. Roth, Avatars for teleconsultation: effects of avatar embodiment techniques on user perception in 3D asymmetric telepresence, *IEEE Transactions on Visualization and Computer Graphics*, 2021, **27**, 4129-4139, doi: 10.1109/TVCG.2021.3106480.

[76] A. M. Villanueva, Z. Liu, Z. Zhu, X. Du, J. Huang, K. A. Peppler, K. Ramani, RobotAR: an augmented reality compatible teleconsulting robotics toolkit for augmented makerspace experiences, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan, 2021, 1-13, doi: 10.1145/3411764.3445726.

[77] D. Saffo, S. Di Bartolomeo, C. Yildirim, C. Dunne, Remote and collaborative virtual reality experiments *via* social VR platforms, *Proceedings of the CHI Conference on Human Factors in Computing Systems,* Yokohama, Japan, 2021, 1-15, doi: 10.1145/3411764.3445426.

[78] B. Marques, S. Silva, A. Rocha, P. Dias, B. S. Santos, Remote asynchronous collaboration in maintenance scenarios using augmented reality and annotations, *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Lisbon, Portugal, March 27-April 1, 2021, 567-568., doi: 10.1109/vrw52623.2021.00166.

[79] D. Fang, H. Xu, X. Yang, M. Bian, An augmented reality-based method for remote collaborative real-time assistance: from a system perspective, *Mobile Networks and Applications*, 2020, **25**, 412-425, doi: 10.1007/s11036-019-01244-4.

[80] P. Wang, X. Bai, M. Billinghurst, S. Zhang, D. Han, M. Sun, Z. Wang, H. Lv, S. Han, Haptic feedback helps me? a VR-SAR remote collaborative system with tangible interaction, *International Journal of Human–Computer Interaction*, 2020, **36**, 1242-1257, doi: 10.1080/10447318.2020.1732140.

[81] I. Fernández del Amo, J. Erkoyuncu, R. Vrabič, R. Frayssinet, C. Vazquez Reynel, R. Roy, Structured authoring for AR-based communication to enhance efficiency in remote diagnosis for complex equipment, *Advanced Engineering Informatics*, 2020, **45**, 101096, doi: 10.1016/j.aei.2020.101096.

[82] S. Kim, G. Lee, M. Billinghurst, W. Huang, The combination of visual communication cues in mixed reality remote collaboration, *Journal on Multimodal User Interfaces*, 2020, **14**, 321-335, doi: 10.1007/s12193-020-00335-x.

[83] P. Wang, X. Bai, M. Billinghurst, S. Zhang, W. He, D. Han, Y. Wang, H. Min, W. Lan, S. Han, Using a head pointer or eye gaze: the effect of gaze on spatial AR remote collaboration for physical tasks, *Interacting with Computers*, 2020, **32**, 153-169.

[84] P. Mohr, S. Mori, T. Langlotz, B. H. Thomas, D. Schmalstieg, D. Kalkofen, Mixed reality light fields for interactive remote assistance, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2020, 1-12, doi: 10.1145/3313831.3376289.

[85] D. Mourtzis, V. Siatras, J. Angelopoulos, Real-time remote maintenance support based on augmented reality (AR), *Applied Sciences*, 2020, **10**, 1855, doi: 10.3390/app10051855.

[86] B. Thoravi Kumaravel, F. Anderson, G. Fitzmaurice, B. Hartmann, T. Grossman, Loki: facilitating remote instruction of physical tasks using bi-directional mixed-reality telepresence, *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, New Orleans, USA, 2019, 161-174, doi: 10.1145/3332165.3347872.

[87] T. Piumsomboon, G. A. Lee, A. Irlitti, B. Ens, B. H. Thomas, M. Billinghurst, On the shoulder of the giant: a multi-scale mixed reality collaboration with 360 video sharing and tangible interaction, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Glasgow Scotland, UK, 2019, 1-17, doi: 10.1145/3290605.3300458.

[88] W. Huang, S. Kim, M. Billinghurst, L. Alem, Sharing hand gesture and sketch cues in remote collaboration, *Journal of Visual Communication and Image Representation*, 2019, **58**, 428-438, doi: 10.1016/j.jvcir.2018.12.010.

[89] S. Kim, G. Lee, W. Huang, H. Kim, W. Woo, M. Billinghurst, Evaluating the combination of visual communication cues for HMD-based mixed reality remote collaboration, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Glasgow Scotland, UK, 2019, 1-13, doi: 10.1145/3290605.3300403.

[90] P. Wang, S. Zhang, X. Bai, M. Billinghurst, L. Zhang, S. Wang, D. Han, H. Lv, Y. Yan, A gesture- and head-based multimodal interaction platform for MR remote collaboration, *The International Journal of Advanced Manufacturing Technology*, 2019, **105**, 3031-3043, doi: 10.1007/s00170-019-04434-2.

[91] P. Wang, S. Zhang, X. Bai, M. Billinghurst, W. He, M. Sun, Y. Chen, H. Lv, H. Ji, 2.5DHANDS: a gesture-based MR remote collaborative platform, *The International Journal of Advanced Manufacturing Technology*, 2019, **102**, 1339-1353, doi: 10.1007/s00170-018-03237-1.

[92] F. Bruno, L. Barbieri, E. Marino, M. Muzzupappa, L. D'Oriano, B. Colacino, An augmented reality tool to detect and annotate design variations in an Industry 4.0 approach, *The International Journal of Advanced Manufacturing Technology*,

2019, **105**, 875-887, doi: 10.1007/s00170-019-04254-4.

[93] I. García-Pereira, C. Portalés, J. Gimeno, S. Casas, A collaborative augmented reality annotation tool for the inspection of prefabricated buildings, *Multimedia Tools and Applications*, 2020, **79**, 6483-6501, doi: 10.1007/s11042-019-08419-x.

[94] K. El Ammari, A. Hammad, Remote interactive collaboration in facilities management using BIM-based mixed reality, *Automation in Construction*, 2019, **107**, 102940, doi: 10.1016/j.autcon.2019.102940.

[95] J. Wang, X. Fan, Y. Zhu, X. Yang, X. Yin, Cross-platform AR annotation for assembly-design communication in pipe outfitting, *The International Journal of Advanced Manufacturing Technology*, 2022, **121**, 4981-4998, doi: 10.1007/s00170-022-09637-8.

[96] K. Kim, M. Billinghurst, G. Bruder, H. B. Duh, G. F. Welch, Revisiting trends in augmented reality research: a review of the 2nd decade of ISMAR (2008–2017), *IEEE Transactions on Visualization and Computer Graphics*, 2018, **24**, 2947-2962, doi: 10.1109/TVCG.2018.2868591.

[97] J. Lazar, J.H. Feng, H. Hochheiser, Research methods in human-computer interaction, *Secnd Edition*, Morgan Kaufmann, Cambridge, 2017, 25-398, ISBN: 978-0-12-805390-4.

[98] G. W. Scurati, F. Ferrise, M. Bertoni, Extended realities and discrete events simulations: a systematic review to define design trade-offs and directions, *Computers in Industry*, 2025, **164**, 104188, doi: 10.1016/j.compind.2024.104188.

[99] R. Hamdani, I. Chihi, Adaptive human-computer interaction for industry 5.0: a novel concept, with comprehensive review and empirical validation, *Computers in Industry*, 2025, **168**, 104268, doi: 10.1016/j.compind.2025.104268.

[100] J. Chen, X. Lu, Y. Du, M. Rejtig, R. Bagley, M. Horn, U. Wilensky, Learning agent-based modeling with LLM companions: experiences of novices and experts using ChatGPT & NetLogo chat, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2024, 1-18., doi: 10.1145/3613904.3642377.

[101] M. Feick, T. Mok, A. Tang, L. Oehlberg, E. Sharlin, Perspective on and re-orientation of physical proxies in object-focused remote collaboration, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Canada, 2018, 1-13, doi: 10.1145/3173574.3173855.

Open Access