NRC·CNRC

NRC Publications Archive
Archives des publications du CNRC

A literature review of AR-based remote guidance tasks with user studies

Lapointe, Jean-François; Molyneaux, Heather; Allili, Mohand Saïd

This publication could be one of several versions: author’s original, accepted manuscript or the publisher’s version. / La version de cette publication peut être l’une des suivantes : la version prépublication de l’auteur, la version acceptée du manuscrit ou la version de l’éditeur.
For the publisher’s version, please access the DOI link below./ Pour consulter la version de l’éditeur, utilisez le lien DOI ci-dessous.

Publisher’s version / Version de l'éditeur:
https://doi.org/10.1007/978-3-030-49698-2_8
Virtual, Augmented and Mixed Reality. Industrial and Everyday Life Applications, Lecture Notes in Computer Science, pp. 111-120, 2020-07-10

NRC Publications Archive Record / Notice des Archives des publications du CNRC :
https://nrc-publications.canada.ca/eng/view/object/?id=bf131768-2e0d-436f-95ec-d390eb731369
https://publications-cnrc.canada.ca/fra/voir/objet/?id=bf131768-2e0d-436f-95ec-d390eb731369

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at
https://nrc-publications.canada.ca/eng/copyright
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L’accès à ce site Web et l’utilisation de son contenu sont assujettis aux conditions présentées dans le site
https://publications-cnrc.canada.ca/fra/droits
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D’UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n’arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research Council Canada
Conseil national de recherches Canada

Canada

A Literature review of AR-based Remote Guidance Tasks with User Studies

Jean-François Lapointe[1][0000-0003-2373-8035], Heather Molyneaux[1][0000-0003-0673-7815] and
Mohand Saïd Allili[2]

[1] National Research Council of Canada, Canada
`{jean-francois.lapointe, heather.molyneaux}@nrc-cnrc.gc.ca`

[2] University of Quebec in Outaouais, Gatineau, Canada
`mohandsaid.allili@uqo.ca`

Abstract. The future of work is increasingly mobile and distributed across space and time. Institutions and individuals are phasing out desktops in favor of laptops, tablets and/or smart phones as much work (assessment, technical support, etc.) is done in the field and not at a desk. There will be a need for systems that support remote collaborations such as remote guidance. Augmented reality (AR) is praised for its ability to show the task at hand within an immersive environment, allowing for spatial clarity and greater efficiency, thereby showing great promise for collaborative and remote guidance tasks; however, there are no systematic reviews of AR based remote guidance systems. This paper reviews the literature describing AR-based remote guidance tasks and discusses the task settings, technical requirements and user groups within the literature, followed by a discussion of further areas of interest for the application of this technology combined with artificial intelligence (AI) algorithms to increase the efficiency of applied tasks.

Keywords: Remote guidance, Augmented Reality, Artificial Intelligence

1 Introduction

The future of work is increasingly mobile and distributed across space and time. Institutions and individuals are phasing out desktops in favor of laptops, tablets, smart phones and/or other mobile digital devices as much work (assessment, technical support etc.) is done in the field and not at a desk. In situations where the assessment, repair or maintenance of a system requires specific expertise, an expert in the field is usually flown in to accomplish the task. This way of doing thing can be costly given the time and money required to travel to the site, thus resulting in reduced productivity.

The advent of mobile augmented reality (AR)-capable digital devices, along with wireless communication networks and artificial intelligence, paves the way to a radical change in work environments. By combining these technologies, we could ask a local novice worker to complete complex tasks under guidance of a remote expert helper who can be located anywhere. This way of doing thing will significantly increase the productivity and/or reduce the cost and response time related to such interventions.
This paper reviews the literature describing AR-based remote guidance tasks and discusses the task settings, technical requirements and user groups within the literature,

followed by a discussion of further areas of interest for AR researchers. It will also discuss possibilities of integrating AI tools such as deep leaning for enhancing the efficiency and accuracy of AR-based remote guidance. AR is a system whereby the real environment is combined with a virtual environment, created digitally, and is interactive in real time. Unlike virtual reality where the environment is entirely rendered, AR enhances the real world by adding virtual information and also could include applications that can remove real objects [1].

The immersive qualities of AR show great promise for increased spatial clarity and greater efficiency for remote guidance tasks; however, we found no systematic review of academic literature examining the use of AR for such tasks. This paper provides a literature review of remote guidance tasks with user studies, followed by information on typical scenarios whereby AR for remote guidance could be used, as well as a discussion on future areas of interest for researchers. This paper finally discuss about various possibilities offered by the combination of AR and AI technologies to create some automation in the process in order to increase the effectiveness of current AR-based remote guidance tasks in concrete scenarios such as those faced by first responders [2, 3, 4, 5].

1.1 Augmented Reality

The concept of AR has been studied for some time now [6, 7, 8] and sits somewhere in the middle of the virtuality continuum [9]. AR is created through the use of physical display devices which can include: fixed devices (tabletop, window); handheld devices (HHD) (phones/tablets) [1]; head-worn devices (HWD) (either optical see through transparent display or video see through opaque display); and projection-based display (PJD) (whereby videos are projected onto objects to be augmented) [10]. Also, other optical devices such as microscopes can benefit from AR [11].

1.2 Remote Guidance

Remote guidance (a.k.a. remote assistance [12], remote coaching [13] or tele-guidance [14]) is a collaboration scenario involving a remote helper guiding in real time a local worker in performing a task on physical objects [15]. Remote guidance can be used by first responders as well as in various fields ranging from medicine, to mining to manufacturing [15]. Remote guidance tasks are typical use cases for AR applications. These tasks can be done via augmented visual instructions with audio either prerecorded [13] or live, coming from a human expert or generated through artificial intelligence [11]. Visual instructions can vary – from directional arrows or lines, shapes, varying colors, virtual hands or other virtual markers for better understanding of the relationship of objects to the environment [13, 15, 16]. The goal is to harness both human and AI capabilities to identify objects and people (for example through object or face recognition) in order to generate virtual content that best fit the real scene and optimize the guidance process.

These types of visual instructions could be applied: to remote guidance for monitoring, assembly [17]; maintenance of equipment in factories [17]; remote guidance for the medical evaluation and intervention on a patient [18]; or remote guidance for disaster damage assessment and fighting for first responders. In the latter case, the technology can be coupled with visual inspection by drones [19] that are remotely guided and extract specific information of interest that can be shared live with the various parties (in-field first responders as well as command and control centers), to maximize the effectiveness and the efficiency of the operations by providing real-time critical information to support the various decision-making processes.

1.3 AR-based Remote Guidance

Several papers already discussed AR-based remote guidance tasks. Some of the early papers to discuss the topic [20, 21, 22, 23] indicate that most of the current systems for remote collaboration (e.g. email as well as audio/video conferencing tools) are designed to support group activities that can be performed without reference to the external spatial environment (e.g. decision making) and that the development of systems to support collaborative physical (a.k.a. spatial workspace collaboration, collaborative physical or remote guidance) tasks has been much slower.

Another indication from Fussell et al. [23] is that information exchanges during collaborative physical tasks generally focus on the identification of target objects, descriptions of actions to be performed on those targets, and confirmation that the actions have been performed successfully. Also, as they speak, collaborators on physical tasks use gestures to clarify or enhance their messages. This paper also describes tools developed to provide remote collaborators the ability to make certain types of gestures by overlaying images, such as a cursor pointer or pen-based drawings, on a live video feed from a workspace. This paper also indicates that collaboration on physical tasks can vary along several dimensions such as the number of participants, temporal dynamics (synchronous vs. asynchronous), type and size of the objects, etc. Like this paper, we will focus here on synchronous collaboration between two remote persons (the local novice worker and the remote expert helper).

For example, Huang and colleagues [15] describe the use of a head-tracked stereoscopic HMD, with sensors and an optical tracker that allows the helper to be immersed in the virtual 3D space of the worker's workspace allowing for two users to see one another's gestures while both working in the field using a hands free system involving a helmet, camera and near-eye display. This user study showed that the use of a 3D immersive interface is helpful for improving user's perception of spatial relation and their sense of co-presence, more particularly for complex tasks.

In another paper [24], Huang and colleagues acknowledge that one of the main issues associated with remote guidance is the loss of common ground. When co-located, collaborators share common ground and are able to constantly use hand gestures to clarify and ground their messages while communicating with each other verbally. However, when collaborators are geographically distributed, such common ground no longer exists, resulting in them not being able to communicate the same way as they do when co-located. As a result, Huang and colleagues focus on the innovation of a shared visual space for gestures in their lab study.

In a separate paper, Fussell and colleagues study a remote guidance task in a case where the remote helper is a mobile robot [23]. In that case, experimental results showed that head movements from the remote robot were effective in the sequential organization of communication. On the other hand, the remote pointing function did not operate ideally within the two reference system of both the instructor and the robot operator.

In [25], Yamashita et al's study has remote collaborators using distributed tabletops activities involving real objects with a technique called "remote lag" to alleviate the problems caused by the invisibility of remote gestures. The invisibility being created by occlusion problems on the tabletops or by remote gestures that were missed when a worker concentrated on the work at hand or his/her attention was directed elsewhere or simply by supposition of visibility. The "remote lag" technique consists in an instant playback of the remote gestures which gives a chance to recover from the missed context of coordination. The results of the experiment presented in [25] suggest that this technique greatly reduces the negative effects of the invisibility problems.

Although remote guidance with pointing (with a laser or a mouse) is an important aspect of remote guidance, research results indicates that projecting hands of the helper supports a much richer set of non-verbal communication and is therefore more effective for remote guidance [26].

Finally, Alem et al. [27] suggest that the following requirements for AR remote guidance system in industry:

1) The need for mobility of the worker by using wearables computers and cameras;
2) The need to allow helpers to guide remotely using their hands.

Based on these requirements, a system called HandsOnVideo was built that uses both standard view and panoramic view cameras to provide both a local and a more global view in order to increase the situational awareness of the users. The system has been used by more than twelve participants for three representative tasks, namely: repairing a photocopy machine; removing a card from a computer motherboard and assembling Lego toys. The worker user interface (UI) in this case used a near-eye display and the overall response from the participants' pool is that the system was quite intuitive and easy to use with no discomfort with the near eye display of the worker system.

2 Combining AR and AI

The basic features characterizing AR systems consist of: 1) analyzing 2D/3D scene to detect positions where virtual content is to be added to the real content, 2) accurate 3D registration of virtual and real content and 3) real-time human-machine interaction [28, 29, 33]. The ultimate goal of these systems is to ensure a seamless blending of virtual and real visual content such that an immersed user perceives it a one continuous and natural environment. Apart this requirement, AR-based remote guidance applications pose additional challenges since their aim is also to guide a remote (novice) technician to perform complex tasks, such as repair or maintenance, sometimes subject to space and time constraints [30, 41].

One promising way to build efficient and automatic systems for AR-based remote guidance is to integrate AI technologies. Indeed, AR systems require basic perception of the visual world and its component elements, estimating the position and orientation of 3D objects of interest and accurately aligning virtual content (e.g., object models, annotations) with the real one [31]. Computer vision and machine learning techniques can provide a viable solution to address these challenges, whereby the built AR systems can integrate modules to automatically analyze and parse the visual environment, generate the appropriate virtual content and align it correctly with the real one.

In the past, several methods have used computer vision techniques to enhance the AR basic features for remote guidance systems. Proposed approaches tried mainly to integrate object detection algorithms in remotely guided maintenance or repair applications using AR-based systems. To track objects on interest in the scene, some methods used markers to provide accurate 3D object position in real time [34]. Markers are visual cues which trigger the display of the virtual information. They are normal images or small objects (e.g., rectangles) that are trained beforehand so that they can be easily recognized in the scene. They can indicate what the user is looking at and where the virtual content should be added. However, they are less adaptive to the environment and are sensitive to occlusions.

Other methods use feature points extracted from images to establish object position by matching these points with a preprocessed database of the objects [35]. For example, Gurevich et al. [29] have proposed the TeleAdvisor system to support remote assistance tasks. It is a hands-free transportable device composed of a video camera and a small projector mounted at the end of a tele-operated robotic arm. The system enables a remote expert to view and interact with a technician workspace, while controlling the point of view. To enable accurate AR projections on the scene, the system uses an active tracking procedure estimating the distance between the surface and the device.

In [30], the authors proposed the ARgitu system for AR-based maintenance of robot arms. To deal with the problem of 3D non-Lambertian surfaces, the system integrates computer-aided design (CAD) object models and conics to build an accurate and robust object detection and tracking module. In addition, the system integrates 3D object training based on the CAD geometric features of each object. It uses also a general easy-to-use authoring tool for developing new (virtual) content for remote guidance in advanced manufacturing industry.

Registration is a necessary step in any interactive AR system. It relates to correctly aligning annotations and virtual content with the real view of the environment [1, 31]. Registration accuracy can be affected by relative motion between the camera and objects, changes of object aspects, object occlusions and changes of lighting conditions. To address this issue, past methods have used mainly markers [29] or feature matching [30] to find the best alignment of virtual and real content. Markers are usually easy to detect and track on images and can ensure speed, stability and robustness of the registration, but it does not adapt to the environment relief. Feature matching consists in visually tracking object features such as color, texture, corners, lines or conics [31]. Since features change in appearance with the point of view, these techniques use complex algorithms, which can decrease the robustness of registration. Finally, most of these methods are not fully automatic, since a human is usually needed to extract features in the first image and to match them with the virtual model.

Recently, deep learning (DL) based methods have shown impressive results for object detection and recognition [34, 35, 36]. Contrarily to traditional machine learning methods that use hand-designed features, DL methods extract features automatically from large amounts of data. In addition, they have robustness to partial occlusion [36].

In [32], the authors proposed DL for recognizing hand actions captured by static RGBD cameras. The method combines recognition of local interactions and global progress of an activity (e.g., assembly, repair, etc.) to feed the user in real-time with actions and context-sensitive prompts while performing the next step of the operation or recovering from errors. The authors use a DL-based segmentation algorithm [33] to localize hands and recognize their local interactions. They also use a probabilistic model for activity state prediction. Finally, Akgul et al. [40] have used DL for accurate and real-time target tracking and detection in AR systems, which showed very promising results.

Another intrinsic part of AR-based remote guidance systems is the virtual content creation service, known as content authoring. It enables the user to create representations of physical objects such as 3D point cloud models, annotated anchor points, buttons, text, animations and virtual objects, which are then added to the real content of the scene. AR-based systems require suitable authoring tools for the development the virtual content and correctly aligning it with the real one. Most of existing methods usually require the use of complex design tools requiring advanced programming [29, 30, 31]. The advent of DL methods can play a key part for virtual content generation. For example, generative neural networks can help in designing realistic 3D and 3D visual content by training on data [37], thus providing a huge potential for AR authoring.

3 Discussion

AR-based remote guidance can be useful in various domains. There are several key components to the use of AR in remote guidance tasks. In general, a good system requires audio communication as well as means for gesture communication (minimally pointing and ideally sketching) and a shared visual space [15, 21, 23, 24, 27]. These requirements place emphasis on the need for environments which have limited background noise, adequate lighting, and a reliable communication link.

Observational studies of physical collaboration suggest that people's speech and action are intricately related to the position and dynamics of objects, other people, and ongoing activities in the environment [23]: therefore, we can already determine that noisy environments are not good candidates for remote guidance unless we can isolate the remote collaborators from that noise.

The same could be said about lighting, i.e. the scene that is to be augmented must be well lit since AR is currently primarily a vision-based technology. Therefore, night or underground operations must provide either appropriate lighting or rely on the use of night-vision systems which themselves rely on image enhancements and/or thermal imaging.

Also, remote guidance requires a communication link between the remote expert helper and the local novice worker. Hence, remote guidance cannot be used in remote areas where no communication link exists or where the latency on the communication link does not allow real-time feedback between both the novice and the expert.

As for the combination of AI and AR, AI technologies such as DL have demonstrated a huge potential to enhance the efficiency and robustness of AR-based remote guidance. Recent advances in DL-based methods including semantic segmentation [33], object detection and recognition [33-36] and image/video annotation [38, 39] offer huge opportunities to build efficient AR-based guidance systems that can enjoy full automation while being adaptive to the environment. These systems can run in real-time while optimizing the guidance process by appropriate augmented content such as sketches, prompts and annotations. Automatic estimation of depth and 3D localization of object boundaries in real-time [33] can also enhance registration accuracy to create immersive experiences of a natural and continuous environment.

Future of AI is very promising where AR systems will have direct benefits from the coming advances. For example, systems can be developed to take into account multiple data modalities (e.g., image, video, sound and text) and spatial and temporal context to optimize the guidance process [41]. Advances in robotics can also enable to help or assist directly novice technicians through a robot arm that performs complex and delicate tasks under the guidance of a remote knowledgeable expert. In addition, systems can employ several cameras to analyze the actions of technician and estimate the progress of a task, while analyzing the object details and state upon which the repair/maintenance task is performed. Semantic segmentation [33] can be used in that regard to identify meaningful regions in an image, such as human hands, buttons or components of an equipment.

Finally, advances of AI can lead to systems enabling self-guidance or remote guidance using robots. Technician workers can often face challenging situations or struggle with equipment malfunctions on sites, while located far from a well-equipped lab and knowledgeable experts. Visual augmentations generated by an onsite or remote computer, which receives input from one or multiple onsite cameras, can be projected to the technician at the physical scene or a see-through device such as smart glasses [42].

4 Conclusion

The literature review on AR-based remote guidance tasks and related user studies indicates that these systems could be used in various fields ranging from medicine to manufacturing to disaster management by first responders. The main usefulness being in the ability for a remote expert helper to guide a local novice worker through the use of the main communications channels (i.e. speech gesture and vision). Within any AR-based remote guidance system, the users must then be provided with a shared visual space (usually a live video feed) and the capability to communicate by speaking, pointing and/or drawing and hand gesturing in order to maximize the usability of the systems. There are specific conditions (good lighting, low acoustic noise, network access) under which using AR for remote guidance would be better suited, according to the work environment as well as the technical requirements needed to establish and maintain a reliable communication link. We also saw that the addition of artificial intelligence (AI) algorithms to these systems can increase the efficiency of applied tasks, and this will likely be an avenue of interest for researchers interested in the use of AR for remote guidance.

References

1. Bonnet, P., Ducher, P., Kubiak, A.: A Brief Introduction to Augmented Reality. In Koelle, M., Linderman, P., Stockinger, T., and Kranz, M (Eds). Human-Computer Interaction with Augmented Reality. Advances in Embedded Interactive Systems Technical Report, vol. 2 (4), 30 pages (2014).
2. Aameer R. W., Sofi, S., Roohie N.: Augmented Reality for Fire & Emergency Services. In: Proceedings of International Conference on Recent Trends in Communications and Computer Networks, pp. 32-41 (2013).
3. Yang, L., Liang, Y., Wu, D., Gault, J.: Train and Equip Firefighters with Cognitive Virtual and Augmented Reality. In: Proceedings of the IEEE 4th International Conference on Collaboration and Internet Computing, pp. 453-459 (2018).
4. Yun, K., Lu, T. Chow, E.: Occluded object reconstruction for first responders with augmented reality glasses using conditional generative adversarial networks. In: Proceedings of SPIE 2018, Volume 10649, 7 pages (2018).
5. Lochhead, I., Hedley, N.: Mixed reality emergency management: bringing virtual evacuation simulations into real-world built environments. International Journal of Digital Earth, 12(2), 190–208 (2019).
6. Caudell, T. P., Mizell, D.W.: Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes. Proceedings of Hawaii International Conference on System Sciences, pp. 659-669 (1992).
7. Azuma, R. T.: A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments, 6(4), 355-385 (1997).
8. Billinghurst, M. Clark, A., Lee, G.: A Survey of Augmented Reality. Foundations and Trends in Human-Computer Interaction, 8(2-3), 73–272 (2014).

9. Milgram, P.: A Taxonomy of Mixed Reality Visual Displays. IEICE Transactions on Information Systems. Vol. E77-D, No. 12, pp. 1321-1329. (1994).
10. Biseria, A., Rao, A.: Human Computer Interface-Augmented Reality. International Journal of Engineering Science and Computing 6(8), 2594-2595 (2016).
11. Chen, P.H.C, Gadepalli, K., MacDonald, R., Liu, Y., Kadowaki, S., Nagpal, K., Kohlberger, T., Dean, J., Corrado, G.S, Hipp, J.D., Mermel, C.H., Stumpe, M.C.: An Augmented Reality Microscope with Real-time Artificial Intelligence Integration for Cancer Diagnosis. Nature Medicine 25, August 2019, 1453-1457 (2019).
12. Schlueter, J.A.: Remote maintenance assistance using real-time augmented reality authoring. Master thesis, Iowa State University, 72 pages (2018).
13. Kim, Y., Hong, S., Kim, G. J.: Augmented Reality based Remote Coaching System. In: Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST 2016), pp. 311-312, Munchen Germany (2016).
14. Yamamoto, T., Otsuki, M., Hideaki, K., H., Suzuki, Y.: Tele-Guidance System to Support Anticipation during Communication. Multimodal Technologies and Interaction, 2(55), 12 pages (2018).
15. Huang, W., Alem, L., Tecchia, F.: HandsIn3D: Supporting Remote Guidance with Immersive Virtual Environments. In: Proceedings of Interact 2013, pp. 70–77, Cape Town, South Africa (2013).
16. Rao, H., Fu, W.-T.: Combining Schematic and Augmented Reality Representations in a Remote Spatial Assistance System. IEEE ISMAR 2013 Workshop on Collaboration in Merging Realities, 6 pages (2013).
17. Hoover, M.: An evaluation of the Microsoft HoloLens for a manufacturing-guided assembly task. Graduate Theses and Dissertations, Iowa State University, 16378, (2018). https://lib.dr.iastate.edu/etd/16378
18. Augestad, K.M., Han, H., Paige, J., Ponsky, T., Schachta, K.M., Dunkin, B., Mellinger, J.: Educational Implications for Surgical Telemonitoring: A Current Review with Recommendations for Future Practice, Policy, and Research. Surgical Endoscopy 31, 3836-3846 (2017).
19. Jones, B., Dillman, K., Tang, R., Tang, A., Sharlin E.,Oehlberg, E., Neustaedter, C., Bateman, S.: Elevating Communication, Collaboration, and Shared Experiences in Mobile Video through Drones. In: Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS 2016), pp. 1123-1135, Brisbane, Australia (2016).
20. Kuzuoka, H.: Spatial workspace collaboration: a SharedView video support system for remote collaboration capability. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'92), pp. 533–540. Monterey, California, USA (1992).
21. Ou, J., Fussell, S.R., Chen, X., Setlock, L.D., Yang, J.: Gestural communication over video stream: supporting multimodal interaction for remote collaborative physical tasks. In: Proceedings of the ACM 5th Conference on Multimodal Interfaces (ICMI'03), pp. 242–249. Vancouver, British Columbia, Canada (2003).
22. Kuzuoka, H., Kosaka, J., Yamazaki, K., Suga, Y., Yamazaki, A., Luff, P., Heath, C.: Mediating dual ecologies. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp. 477–486. Chicago, Illinois, USA (2004).
23. Fussell, S. R., Setlock, L.D., Yang, J., Ou, J., Mauer, E., Kramer, A.D.L.: Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks. Human-Computer Interaction 19, 273-309, (2004).
24. Huang, W, Alem, L., Tecchia, F., Been-Lirn Duh, H.: Augmented 3D hands: a gesture-based mixed reality system for distributed collaboration. Journal of Multimodal User Interfaces 12(2), 77-89 (2018).

25. Yamashita N, Kaji K, Kuzuoka H, Hirata K (2011) Improving visibility of remote gestures in distributed tabletop collaboration. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW 2011), pp. 95–104. Hangzhou, China (2011).
26. Kirk, S. D., Fraser, D. S.: Comparing Remote gesture Technologies for Supporting Collaborative Physical Tasks. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'06), pp. 1191-1200. Montréal, Québec, Canada (2006).
27. Alem, L., Tecchia, F., Huang, W.: HandsOnVideo: towards a gesture based mobile AR system for remote collaboration. In:Alem L., Huang W.(eds) Recent trends of mobile collaborative augmented reality. Springer, New York, pp 127–138. (2011).
28. Herskovitz, J., Ofek, E., Lasecki, W.S., Fourney A.: Opportunities for In-Home Augmented Reality Guidance. In: Proceedings of the ACM Conference on Human Factors in Computing Systems CHI Extended Abstracts, pp. 1-6. Glasgow, UK (2019).
29. Gurevich, P., Lanir, J., Cohen, B., Stone, R.: TeleAdvisor: A Versatile Augmented Reality Tool for Remote Assistance. In: Proceedings of the ACM Conference on Human Factors in Computing Systems CHI, pp. 619-622, Austin, Texas, USA (2012).
30. Zubizarreta, J., Aguinaga, I., Amundarain, A., A framework for augmented reality guidance in industry. International Journal of Advanced Manufacturing Technology, 102(9-12), 4095–4108 (2019).
31. Didier, J.-Y., Rousse, D., Mallem, M., Otmane, S., Naudet, S., Pham, Q.-C., Bourgeois, S., Mégard, C., Leroux, C., Hocquard, A.: AMRA: Augmented Reality Assistance for Train Maintenance Tasks. In: Proceedings of the 4th ACM/IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2005): Workshop Industrial Augmented Reality, pp. 1-10, Vienna, Austria (2005).
32. Schröder, M., Ritter, H.J.: Deep learning for action recognition in augmented reality assistance systems. In: Proceedings of SIGGRAPH, pp. 1-2, Los Angeles, California, USA (2017).
33. Hu, R., Dollár, P., He, K., Darrell, T., Girshick, R.B.: Learning to Segment Every Thing. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), pp. 4233-4241, Salt Lake City, Utah, USA (2018).
34. Ren, S.He, K., Girshick, R., Sun, J.:Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: IEEE Transactions on Pattern Analysis & Machine Intelligence, 39(6), 1137-1149 (2017).
35. Redmon, J., Divvala, S., Girshick, R., Farhadi A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779–788, (2016).
36. Lin, T.-Y., Goyal, P., Girshick, R.B, He, H., Dollar, P.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 2999-3007, Venice, Italy (2017).
37. Ro, H., Park, Y.-J., Byun. J., Han, T.-D.: Display methods of projection augmented reality based on deep learning pose estimation. In: Proceedings of the ACM SIGGRAPH'19 Posters, pp. 1-2, Los Angeles, California, USA (2019).
38. Laib, L., Allili, M.-S., Ait-Aoudia, S.: A probabilistic topic model for event-based image classification and multi-label annotation. Signal Processing: Image Communication, 76: 283-294 (2019).
39. Ouyed, O., Allili, M.-S.: Feature weighting for multinomial kernel logistic regression and application to action recognition. Neurocomputing, 275, pp. 1752-1768 (2018).
40. Akgul, O., Penekli, H.I., Genc, Y.: Applying Deep Learning in Augmented Reality Tracking. In: Proceedings of the 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), pp. 47-54, Naples, Italy (2016).

41. Zhu, J., Ong, S. K., Nee, A. Y. C.: An authorable context-aware augmented reality system to assist the maintenance technicians. International Journal of Advanced Manufacturing Technology, 66, 1699-1714 (2013).
42. Essig, K. Strenge, B., Schack, T.: ADAMAAS: Towards Smart Glasses for Mobile and Personalized Action Assistance. In: Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments (PETRA), pp. 1-4. Corfu, Greece (2016).