

Usage and Effect of Eye Tracking in Remote Guidance

CHUN XIAO and WEIDONG HUANG, Faculty of Transdisciplinary Innovation, University of Technology Sydney, Australia

MARK BILLINGHURST, School of ITMS, University of South Australia, Australia

Eye gaze information is an important and effective input modality for human-human communication. As eye tracker is becoming more affordable for remote collaboration, a systematic literature review about eye-tracking supported collaboration systems on physical works will improve the collaboration and benefit system design in the future. Toward this end, we review publications since year 2000 and categorize the reported prototypes and systems with respect to eye gaze functionality, eye-tracked subjects, physical task types and gaze visualisations, demonstrating an overview of the usage and effect of eye tracking in remote guidance over time. We identify the usage of eye tracking in remote guidance as gaze behavior and intention indicator, fast and effective input modality, physical referential pointer and communication cues for social presence. Finally we analyze and discuss the effect and limitation of eye tracking in remote guidance with respect to different gaze visualisations in a variety of applied scenarios. The technical and social challenges identified will improve collaborations under remote guidance and benefit eye-tracking supported collaboration system design in the future.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; Interaction paradigms; Interactive systems and tools.

Additional Key Words and Phrases: eye tracking, remote guidance, gaze visualisation, physical work, collaborative work

ACM Reference Format:

Chun Xiao, Weidong Huang, and Mark Billinghurst. 2020. Usage and Effect of Eye Tracking in Remote Guidance. In *OzCHI '20: 32nd Australian Conference on Human-Computer Interaction (HCI), December 02–04, 2020, Online, Australia*. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/1122445.1122456>

1 INTRODUCTION

Working from home is one of the main themes of 2020, and the increase in remote collaboration could have globally significant socioeconomic impact over the long term. In remote collaboration on physical tasks, a local worker can perform physical object manipulations guided by a remote expert [24], saving time and money. However, when compared to face-to-face collaboration, there are challenges in remote collaboration, which includes monitoring the distant workspace, building up communication ground, sharing a referential system, and understanding the remote activities and perceptions of the collaborators via the communication channels available [32].

Eye-contact is an important form of nonverbal communication, and plays an important role in face-to-face conversation. However eye-contact was often missing due to hardware limitations or availability in earlier video-mediated remote communication [4, 26]. Recently eye-tracking technology has begun to be explored in the field of computer supported cooperative work (CSCW) with different configurations, and the effectiveness of introducing eye tracking

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

© 2020 Association for Computing Machinery.

Manuscript submitted to ACM

seems promising. In human-human remote collaboration, eye gaze is an instant and effortless input modality in addition to touch, speech, and hand and other body postures and gestures [27]. Eye gaze, when visible in a shared view or interface during remote collaboration, can present valuable contextual cues about the current visual attention of one participant. When gaze is shared with the other collaborator, it can lead to improved joint attention. The joint attention usually converges to a physical target in a collaborative physical work, so gaze can function as a quick and precise pointing gesture, and is very effective when the focus of attention is complicated to describe verbally or by other ways [6, 35]. In collaborative physical works under remote guidance, gaze is most powerful when used in combination with other more explicit input modalities [27] such as hand gesture [24, 25, 28, 29, 56]. Gaze has been used as a pointer for referring and annotating functions, as many other pointer modalities [1, 9, 20, 30, 49, 52, 57] and annotation systems [17, 21, 41, 45] do. In addition to pointing, research has been undertaken to provide gaze cues in augmented, mixed and virtual reality (AR, MR or VR) for improved immersive collaboration experiences [5, 46, 47, 50].

This leads to our motivation to provide an overview on the current research in gaze usage and effect in remote collaborative systems. We believe that the technical and social challenges identified will improve remote guided collaborations and benefit developing eye-tracking supported collaboration systems in the future.

The main contributions of this paper are listed as follows:

- Firstly, we categorize the prototypes and systems according to eye tracking functionality, eye-tracked subjects, physical task types and gaze visualisations, which lead to an overview of the usage and effect of eye tracking in remote guidance over time.
- Secondly, we summarize the usage of eye tracking in remote guidance as gaze behavior and intention indicator, fast and effective input modality, physical referential pointer and communication cues for social presence.
- In addition to the usage, the effect of eye tracking in remote guidance is also analyzed and discussed based on different gaze visualisations in a variety of applied scenarios.

2 METHOD OF PUBLICATION COLLECTION AND SELECTION

We collected papers from three main sources: the IEEE/IEL electronic library, the ACM digital library and the Springer-Link full text databases. Publications before 2000 were not generally considered, because research about eye-tracking supported collaborations in earlier years was quite restricted due to hardware conditions. We focused the literature review on physical tasks. A physical work or task refers to an object-oriented manipulation work, and so it could be a task with a real workspace scene, or a simulated scene shown on a screen or in a virtual environment, where the participants needed to take operational actions on an object target. Most importantly, the task needed to involve eye tracking.

A group of search keywords and terms were used to search the full texts in each database, which were “*eye tracking*”, “*remote*”, and “*collaborative work*” or “*collaborative task*”. We didn’t set additional search restrictions in order to reach a broad coverage of related publications as much as possible, although we focused on remote collaboration of physical works. By selecting the top relevant 100 publications rated by each database, we collected 165 publications of conference papers, journal articles, and book chapters. The left bar chart of Fig. 1 shows the number of publications collected from each databases. The right word cloud plot of Fig. 1 demonstrates the most frequent keywords and terms in the publications collected, the bigger the font size of a keyword, the higher the frequency it occurs. We noticed that there were quite a number of publications studying mechanisms of gaze behavior, as well as gaze visualisation in collaborative

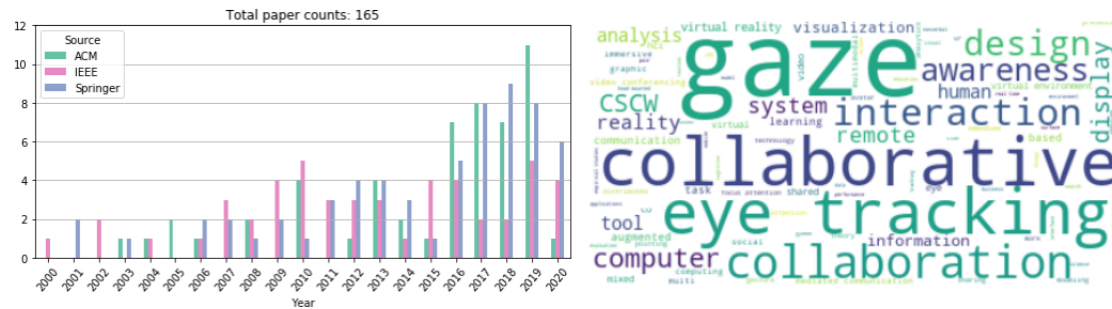


Fig. 1. Number of publications collected from the three databases (left) and the most frequent keywords and terms in the collection (right).

search tasks, noting that we won't discuss all of them in details, instead we mention and discuss the findings reported as reference in the following sections.

It is seen that since 2016 there has been an overall increase in eye-tracking related research publications about remote collaborative works. From the publication collection we also identify that the *ACM Conference series on Human Factors in Computing Systems (CHI)* is the most significant source of publisher and venue for this survey. Other publisher and venue sources include but not limited to *ACM Conference on Computer Supported Cooperative Work and Social Computing Companion (CSCW)*, *ACM Symposium on Eye Tracking Research&Applications (ETRA)*, *International Conference on Multimodal Interfaces (ICMI)*, *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, *IEEE Virtual Reality conference (VR)*, *IEEE transactions*, and a range of Springer journals.

3 AN OVERVIEW OF EYE-TRACKING SUPPORTED REMOTE COLLABORATION PROTOTYPES AND SYSTEMS ON PHYSICAL TASKS

We categorize the prototypes and systems from different perspectives to present an overview of eye-tracking supported remote collaboration prototypes and systems on physical tasks.

3.1 The key role eye tracking playing in the prototypes and systems reviewed

The key role that eye tracking plays in remote collaborative physical works is to identify focus of attention (FOA). The joint attention guided by one participant's FOA, aiming to provide visual referencing for object manipulation or common understanding purpose, as studied by Fussell et al. [19], is key for remote collaborative physical works. Research suggests that gaze is an active and effective trigger of FOA. Specifically, researchers compared the effect of remote guidance under various conditions with gaze and/or other indicators like conversation [3, 13, 23], head direction [54, 55], hand gesture [6, 7, 23], facial expression [10], cursor pointer [22], or a combination of several cues [6, 7, 23]. This is measured using different metrics such as task completion time, number of mistakes, number of words or phrases used in conversation, accompanied with subjective user questionnaires. The introducing of eye tracking generally improves the performance of experiments under certain conditions when compared to a set of different baseline conditions as reported, but eye tracking also shows limitations and disadvantages. This issue will be discussed in section 5.2.

In addition to supporting object manipulation, eye-tracking techniques were applied in virtual environments [5, 15, 50] as well as in augmented, mixed or virtual realities [6, 46, 47] with avatars as representation of participants. This was often done to support the same social interactions between participants as in face-to-face collaborations, which leads to more technical challenges.

3.2 Collaborative tasks

We mainly reviewed publications reporting collaborative object manipulation tasks, but we also found that some prototypes were built up based on simulated tasks visible on a 2D screen or in virtual environments. A typical task studied was a puzzle assembly or block building task, where essential interactions such as actions of referring target and location, concept understanding, or communication expressions, can be very well investigated.

3.3 Eye-tracked subjects

When considering the subjects who were eye-tracked in collaborative physical works under remote guidance, there were three use cases in which eye tracking could be applied: the remote helper only, the local worker only, or both. Based on different research scenarios the systems or experiments were designed accordingly to implement eye tracking on different subjects. In general practice, gaze might indicate a status that someone is monitoring, which seems to happen more often with helpers when monitoring the task status and actions of workers. Physically, eye tracking, together with head position and other body motions and gestures, could work as a FOA indicator or a pointing gesture to formulate messages, which might happen to both helpers and workers during the collaboration [19].

3.3.1 Eye tracking of the remote helper. Quite a few prototypes and systems, when conducting real scene physical tasks under guidance, only applied eye tracking on the remote helpers. These include teleoperation systems in which eye tracking is applied only to the human operators, and eye gaze functions as an operation trigger to control robot workers. Eye tracking of remote helpers is visible in the shared view, display, or the physical workspace, except in early exploratory works due to hardware limitation. The helper's gaze, when overlaid on the worker's view, can successfully guide the FOA of the local worker. For example, Higuch et al. [23] used the gaze cues of a remote helper to guide local physical tasks such as block construction and showroom assembly. Their work also attempted to combine eye fixations and hand gestures with verbal cues. Their results suggested that the helper's gaze is a precise pointing gesture faster than hand gesture.

3.3.2 Eye tracking of the local worker. However, there are few prototypes and systems in which the local workers were eye-tracked but not the remote helpers. The exploratory experiment of Fussel et al. [18] indicated that "*the performance with a head mounted camera (HMC) with eye gaze information from the worker's side was not significantly different than using audio only, and performance was significantly worse than with a fixed camera showing the entire workspace*". Barakonyi et al. [7] introduced their remote AR videoconferencing system where the local student selected an object by eye gaze, and the remote helper (teacher) perceived the target and offered help by explaining the selected target. Later, the investigation of Gupta and Billinghurst [10, 22] demonstrated that the physical task completion time was reduced, and user interactive experience was improved under the condition that the worker was eye-tracked and monitored by a helper, and a mouse pointer representing the helper's referring target was visible in the worker's view, when compared to the condition without the pointer nor eye tracking. It is noticed that in Gupta and Billinghurst's work [10, 22], the worker could observe the view of workspace from the head-mounted display, while the worker could monitor the helper's head and upper body in the work of Fussel et al. [18]. The eye tracking

3.3.3 Dual eye-tracking. Researchers have also investigated gaze coordination mechanisms to detect joint attention. Due to technical and hardware availability issues, many of the dual eye-tracking supported prototypes were set up based on simulated collaborative physical tasks or in virtual environments. Until recent several prototypes were designed to reproduce the participants' gaze in AR, VR and MR for physical collaboration in a real scene. An example is MiniMe [47], which used eye tracking in AR for the local worker and in VR for the remote helper, and gaze cues were visualized as virtual ray-cast. Their study demonstrated that dual eye tracking not only contributes to accomplishing the physical task itself as a physical pointer for physical object manipulation, but also builds up the communicative channel in addition to body gestures, as well as providing a monitor channel for the third party. In particular, dual eye tracking significantly improves the social interactive experience of human participants just as in side-by-side collaboration.

3.4 Visibility and representation of eye gaze

Visible gaze from collaborators explicitly helps to guide the attention or awareness thus establishes the joint attention, which is a critical component for collaboration. With different system design and applied environment, a variety of visualisations are demonstrated in different prototypes and systems.

3.4.1 Gaze visible in 2-dimensional view. We find that gaze cursor has been the most commonly applied visualisation so far, with many forms such as a circle or eye-shaped icon visible on a 2D scene screen. In addition to gaze cursors, Chetwood et al. [13] used hotspots to represent the gaze of a master over time in a simulated laparoscopic surgical equipment to guide surgeon training. The trainee's gaze was visualized the same way and monitored by the master. Higuch et al. [23] visualized the helper's gaze in a combined way for an object manipulation task, where the gaze direction was represented as a piece of highlighted arc when the gaze target was out of the limited HMD field of view of the worker, and the exact gaze focus in the worker's view was visualized as a green square cursor. Li et al. [33] visualized gaze as gaze trail or zoom focus for a simulated collaborative puzzle assembly task and a bomb diffusion task.

During a joint spatial search task, D'Angelo et al. [16] studied and compared using a heatmap, gaze path or shared area as gaze visualisations for a joint on-screen pattern search task. They found that the collaborative visual search task was completed in less time when gaze was visualized as a shared area or gaze path. Zhang et al. [58] studied four gaze visualisations for a co-located map search task and reported that less explicit gaze visualisation was preferred, where the gaze path (trajectory) should be avoid, as high visibility of gaze indicators might cause more distraction. These results also suggest that displaying gaze indicators generally improves the collaboration efficiency and establishes joint attention, but the efficiency of visible gaze vary with different gaze visualisations. Please note that Zhang's work is a co-located collaboration, as their findings are valuable for remote collaborations as well. The investigations above lead to gaze visualisation and eye-tracking data processing and modelling studies.

3.4.2 Gaze visible as a projected pointer in physical workspace. In some studies augmented eye gaze has been directly projected into a relative small workspace. Akkil et al. [3, 4] projected the helper's gaze as a torch/cursor into the worker's physical workspace to guide the building of blocks, as Higuch et al. [23] and Wang et al. [54, 55] did. In Higuch's work, the projected gaze was visualized as gaze path that kept a small piece of gaze travel history, which was quite informative to instruct an object placement action from one place to the destination. Rheden et al. [53] also projected the eye gaze of the helper into a 6*6 grid whiteboard to guide the worker in drawing a shape by connecting grid edges with lines. It should be noted that Rheden's work is a co-located collaboration, however their findings can be of great benefit to remote mode as a reference.

Compared to the case that a worker has to wear a head-mounted device or has to observe a display or a screen showing instructions from the helper, this relatively free-style work mode is more similar to a co-located one. The projected gaze in the real workspace strongly improves the collaboration efficiency.

3.4.3 Gaze as a cursor or ray-cast in VE, AR, VR and MR. Gaze visualisation and transformation becomes more challenging in virtual environments. Despite the technical issues, it is reported that the selection criteria of participants is very critical. Unfortunately candidates with a big head or wide distance between the eyes, or are wearing glasses may not be able to use eye-tracking in a head-mounted display. This results in relative fewer investigations related to eye-tracking in AR, VR or MR.

In virtual environments, Duchowski et al. [15], Steptoe et al [50, 51], Andrist et al. [5] made gaze visible as lightspot or virtual line to better monitor the task status and referenced objects for the collaborator. Piumsomboon et al. [46] tracked gaze of both AR and VR users in a collaborative spatial object manipulation task. Gaze was visualized as a virtual ray-cast, and when the users looked at an object longer together the color of the object changed. Piumsomboon et al. [47] augmented the gaze of both the remote helper and the local worker, where both avatars were shown side-by-side in a mixed reality environment, and gaze was visualized as ray-cast to instruct object manipulations in the real environment.

3.4.4 Without gaze visualisation. Sometimes gaze is not always or can not be visualized explicitly under certain circumstances. The AR videoconferencing interface of Barakonyi et al. [7] set the eye cursor hidden by default, which could be selectively turned on by a keyboard shortcut to avoid meaningless gaze cues for object selection. The ThirdEye developed by Otsuki et al. [40] guided attention in videoconferencing, where the gaze FOA was not explicitly annotated except an anchored eye-shaped button showing eye movements. As demonstrated in their work, the ThirdEye successfully led the attention.

4 USAGE OF EYE TRACKING IN REMOTE GUIDANCE

4.1 Gaze cues as behavior indicator to reveal interaction mechanism

Chanel et al. [12] found that eye-movement coupling is a very important indicator when assessing collaborative interactions regarding *convergence and emotion management*, where convergence refers to action synchrony, mutual understanding, conceptual convergence, emotional convergence, symmetry in roles and responsibilities, and emotion management refers to the emotion awareness of collaborative participants. Rae et al. [48] also mentioned that using gaze in designating or monitoring are the two classes of gaze-based human interaction practices. Fussell et al. [18] investigated the process during a remote guided physical collaboration in which the helper was eye-tracked. Gaze cues were found to be used to *monitor people's actions, establish joint focus of attention, and formulate messages as a pointing gesture*.

4.2 Gaze cues as intention indicator to predict actions

Identifying focus of attention (FOA) is a critical task in remote guided collaboration. The joint attention guided by FOA of one participant strongly supports interactions such as visual referencing for grounding, disambiguation or object manipulation.

Ou et al. [42, 44] identified that the helper's FOA areas were the workspace, the jigsaw piece hub and target puzzle by monitoring the helper's eye movement. Akkil and Isokoski [2] examined the effect of gaze augmentation in a video recorded from a HMC in a simulated driving task, and their results suggested that videos with augmented gaze, along

with the contextual information available, can efficiently trigger the viewers' action to predict the observed people's intention. Higuch et al. [23] also reported that the worker could predict the intention of the helper when the helper's gaze was visible even without speech given a specific context. Moreover, Ou et al. [43] proposed a Markov model to predict the helper's FOA. In addition to Markov model, support vector machines (SVMs) were also applied to predict intention of actions in computer gaming recently [38, 39]. It is quite promising for smart collaborative system design given reliable intention prediction.

4.3 Gaze as a fast non-verbal input modality

The experiments reported in [4, 8, 10, 11, 14, 22, 23] compared conditions when gaze was visible or invisible during remote collaboration. The results suggest that visible gaze improves the collaborative efficiency, as *"eye fixations visualized by both a projector and a HMD show a fast and precise pointing capability over hand gestures"* [23], which is a consistent finding that shared gaze significantly reduced completion time in a face-to-face collaborative construction task, as Huang and Shi [34] reported recently.

Chetwood et al. [13] also compared verbal-based and gaze-based instructions in collaborative surgeon training. They found that eye gaze is especially useful in collaboration between clinicians from different countries or background, given the different languages and physical conditions where clinicians must wear masks make verbal communication inapplicable during the collaboration. Similarly, Kwok et al. [31] indicated that eye gaze is more advantageous in terms of implicitly carrying information on the focus of the surgeon's attention than other input modalities in collaborative surgery.

These investigations reveal that gaze information is a very important input modality supplementing body gesture and speech in collaborative physical tasks under remote guidance.

4.4 Improving co-presence by reproducing eye movement of avatars in virtual environment

As people's gazing practices and the visibility of these to others are important resources in social interaction [48], it is important to reproduce gaze to improve immersion of virtual reality. Virtual systems build in [5, 6, 15, 37, 47, 51, 51], all tried to reproduce the users' gaze in avatars in the virtual environments. The augmented gaze of an avatar not only worked as pointing gesture in object manipulation tasks, but also established a communication channel in augmented, virtual, and mixed realities.

5 DISCUSSION

5.1 Effect of gaze visualisation

From the gaze visualisations presented above, several issues regarding the effect of gaze visualisation can be observed.

- Gaze visualisations vary with respect to different types of tasks and display conditions. Basically, if the collaborative task is monitored and presented on a screen, a gaze cursor is the common choice with quite a few variations. It could be investigated if the mutual visible gaze pattern (the gaze cursor is only visible to the other side) or mirror visible gaze pattern (both gaze cursors are visible in the shared view) can work equivalently in dual eye tracking systems, as the mirror visible gaze pattern may cause distractions. For physical collaboration under remote guidance, a projected gaze pointer in the physical workspace is feasible and improves not only the collaboration efficiency but also the social co-presence. However, the projected gaze might be hard to track in a large physical workspace, gaze simulation via avatars seem to be a good solution. With the development

of hardware, delicate algorithms to capture the behavior of human gaze will contribute to the design of more advanced systems for more complicated tasks.

- The visible gaze in the physical workspace is supposed to improve the collaboration, as the worker won't need to pay additional attention to a shared view screen. However, distractions when showing non-filtered gaze or a personal feeling like "*being watched*" were reported when gaze is visible in the physical workspace [4, 23, 58].
- Visualisation as a gaze path could be applied when moving an object from one location to another. This is supposed to guide object moving actions well, although a gaze path was reported to cause distractions and should be avoid in a collaborative search task [58].

5.2 Effect of gaze pointer

As mentioned above, gaze as a gaze pointer or a referential gesture is very effective when the referenced object is quite complicated to be described verbally or by other input modalities, however it is most powerful when used in combination with other input modalities such as conversation and body gestures. Gaze information alone may not be as effective as a mouse.

In earlier work of Bauer et al. [9], it was found that the pointer/annotation working alone without speech cues didn't help much in communication. Müller et al. [36] also indicated that gaze transfer can help establish common ground in remote cooperation. However, gaze may lead to ambiguity when conveying spatial information, so that a conventional referential device such as a mouse could outperform the gaze in this case. Higuch et al. [23] also mentioned that gaze cues for providing explicit instruction purposes worked only when they were combined with speech. Akkil and Isokoski [3] reported a similar case where the gaze pointer may need more verbal instructions to accomplish a physical task under remote guidance. Consistently, in remote collaborative physical tasks applying augmented, mixed, or virtual reality technology, the gaze cue was reported to work together with other gesture cues better than to work alone, based on a verbal condition. The gaze cue alone worked better when referring to a complicated spatial layout and self-location awareness than the combined cues [6]. However, when comparing cursor pointer, head pointer and eye gaze as pointer, Wang et al. [54] indicated that the head pointer and eye gaze led to better collaboration quality, and the head pointer could be considered as a low-cost proxy of the eye gaze pointer [55].

5.3 Limitation of this work

The systems and prototypes reviewed in this work are limited to physical works, which do not cover many collaboration systems like e-learning, e-writing, pair programming, online shopping, computer gaming, and etc. Therefore the usage and effect of systems discussed is specified within this restricted scope.

Another limitation could be the literature collecting process. as the selected keywords and terms might have only covered a subset of all relevant keywords. The search in full texts might have also caused additional workload of literature review, as keywords in full texts are not always related to the work conducted and reported in the publications.

6 CONCLUSION

We briefly reviewed eye-tracking supported prototypes and systems for remote collaboration on physical works over the last two decades, and presented an overview of usage and effect of eye tracking in remote guidance over time. The prototypes and systems are categorized and analyzed in terms of eye tracking functionality, eye-tracked subjects, physical task types and gaze visualisations, where the usage of eye tracking in remote guidance is summarized and categorized as gaze behavior and intention indicator, fast and effective input modality, physical referential pointer

and social presence cues. The effect and limitation of eye tracking in remote guidance is analyzed and discussed with respect to the usage, which will lead to a better understanding of eye tracking functionality in remote guidance, and benefit remote collaboration system design in the future.

ACKNOWLEDGMENTS

This research was funded fully by the Australian Government through the Australian Research Council.

REFERENCES

- [1] Matt Adcock and Chris Gunn. 2015. Using Projected Light for Mobile Remote Guidance. *Computer Supported Cooperative Work (CSCW)* 24 (2015), 591–611. <https://doi.org/10.1007/s10606-015-9237-2>
- [2] Deepak Akkil and Poika Isokoski. 2016. Gaze Augmentation in Egocentric Video Improves Awareness of Intention. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). ACM, New York, NY, USA, 1573–1584. <https://doi.org/10.1145/2858036.2858127>
- [3] Deepak Akkil and Poika Isokoski. 2019. Comparison of Gaze and Mouse Pointers for Video-based Collaborative Physical Task. *Interacting with Computers* 30, Article 26 (02 2019), 19 pages. <https://doi.org/10.1093/iwc/iwy026>
- [4] Deepak Akkil, Jobin Mathew James, Poika Isokoski, and Jari Kangas. 2016. GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI EA '16*). ACM, New York, NY, USA, 1151–1158. <https://doi.org/10.1145/2851581.2892459>
- [5] Sean Andrist, Michael Gleicher, and Bilge Mutlu. 2017. Looking Coordinated: Bidirectional Gaze Mechanisms for Collaborative Interaction with Virtual Characters. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). ACM, New York, NY, USA, 2571–2582. <https://doi.org/10.1145/3025453.3026033>
- [6] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). ACM, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376550>
- [7] Istvan Barakonyi, Helmut Prendinger, Dieter Schmalstieg, and Mitsuru Ishizuka. 2007. Cascading Hand and Eye Movement for Augmented Reality Videoconferencing. In *2007 IEEE Symposium on 3D User Interfaces*. IEEE, USA, 71–78. <https://doi.org/10.1109/3DUI.2007.340777>
- [8] Ellen Gurman Bard, Robin L. Hill, Mary Ellen Foster, and Manabu Arai. 2014. Tuning accessibility of referring expressions in situated dialogue. *Language, Cognition and Neuroscience* 29, 8 (2014), 928–949. <https://doi.org/10.1080/23273798.2014.895845>
- [9] Martin Bauer, Gerd Kortuem, and Zary Segall. 1999. “Where Are You Pointing At?” A Study of Remote Collaboration in a Wearable Videoconference System. In *Proceedings of the 3rd IEEE International Symposium on Wearable Computers (ISWC '99)*. IEEE, USA, 151–158. <https://doi.org/10.1109/ISWC.1999.806696>
- [10] Mark Billinghurst, Kunal Gupta, Masai Katsutoshi, Youngho Lee, Gun A. Lee, Kai Kunze, and Maki Sugimoto. 2016. *Is It in Your Eyes? Explorations in Using Gaze Cues for Remote Collaboration*. Springer, Switzerland, 177–199. https://doi.org/10.1007/978-3-319-45853-3_9
- [11] Susan Brennan, Xin Chen, Christopher Dickinson, Mark B. Neider, and Gregory Zelinsky. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition* 106 (04 2008), 1465–77. <https://doi.org/10.1016/j.cognition.2007.05.012>
- [12] Guillaume Chanel, Mireille Bétrancourt, Thierry Pun, Donato Cereghetti, and Gaëlle Molinari. 2013. Assessment of Computer-Supported Collaborative Processes Using Interpersonal Physiological and Eye-Movement Coupling. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*. IEEE, USA, 116–122.
- [13] Andrew Chetwood, Ka-Wai Kwok, Loi-Wah Sun, George Mylonas, James Clark, Ara Darzi, and Guang-Zhong Yang. 2012. Collaborative eye tracking: A potential training tool in laparoscopic surgery. *Surgical endoscopy* 26 (01 2012), 2003–2009. <https://doi.org/10.1007/s00464-011-2143-x>
- [14] Sarah D’Angelo and Darren Gergle. 2016. Gazed and Confused: Understanding and Designing Shared Gaze for Remote Collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2492–2496. <https://doi.org/10.1145/2858036.2858499>
- [15] Andrew T. Duchowski, Nathan Cournia, Brian Cumming, Daniel McCallum, Anand Gramopadhye, Joel Greenstein, Sajay Sadasivan, and Richard A. Tyrrell. 2004. Visual Deictic Reference in a Collaborative Virtual Environment. In *Proceedings of the 2004 Symposium on Eye Tracking Research & Applications* (San Antonio, Texas) (*ETRA '04*). ACM, New York, NY, USA, 35–40. <https://doi.org/10.1145/968363.968369>
- [16] Sarah D’Angelo and Darren Gergle. 2018. An Eye For Design: Gaze Visualizations for Remote Collaborative Work. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). ACM, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3173923>
- [17] Omid Fakourfar, Kevin Ta, Richard Tang, Scott Bateman, and Anthony Tang. 2016. Stabilized Annotations for Mobile Remote Assistance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). ACM, New York, NY, USA, 1548–1560. <https://doi.org/10.1145/2858036.2858171>

- [18] Susan R. Fussell, Leslie D. Setlock, and Robert E. Kraut. 2003. Effects of Head-Mounted and Scene-Oriented Video Systems on Remote Collaboration on Physical Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (*CHI '03*). ACM, New York, NY, USA, 513–520. <https://doi.org/10.1145/642611.642701>
- [19] Susan R. Fussell, Leslie D. Setlock, and Elizabeth M. Parker. 2003. Where do Helpers Look? Gaze Targets During Collaborative Physical Tasks. In *Extended abstracts of the 2003 Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (*CHI '03*). ACM, New York, NY, USA, 768–769. <https://doi.org/10.1145/765891.765980>
- [20] Susan R. Fussell, Leslie D. Setlock, Elizabeth M. Parker, and Jie Yang. 2003. Assessing the Value of a Cursor Pointing Device for Remote Collaboration on Physical Tasks. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (*CHI EA '03*). ACM, New York, NY, USA, 788–789. <https://doi.org/10.1145/765891.765992>
- [21] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. World-Stabilized Annotations and Virtual Scene Navigation for Remote Collaboration. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) (*UIST '14*). ACM, New York, NY, USA, 449–459. <https://doi.org/10.1145/2642918.2647372>
- [22] Kunal Gupta, Gun A Lee, and Mark Billinghurst. 2016. Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration. *IEEE Transactions on Visualization and Computer Graphics* 22, 11 (2016), 2413–2422. <https://doi.org/10.1109/TVCG.2016.2593778>
- [23] Keita Higuchi, Ryo Yonetani, and Yoichi Sato. 2016. Can Eye Help You? Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). ACM, New York, NY, USA, 5180–5190. <https://doi.org/10.1145/2858036.2858438>
- [24] Weidong Huang and Leila Alem. 2013. Gesturing in the Air: Supporting Full Mobility in Remote Collaboration on Physical Tasks. *Journal of Universal Computer Science* 19 (01 2013), 1158–1174.
- [25] Weidong Huang, Leila Alem, Franco Tecchia, and Henry Duh. 2017. Augmented 3D hands: a gesture-based mixed reality system for distributed collaboration. *Journal on Multimodal User Interfaces* 12 (11 2017), 77–89. <https://doi.org/10.1007/s12193-017-0250-2>
- [26] Tomoko Imai, Dairoku Sekiguchi, Masahiko Inami, Naoki Kwakami, and Susumu Tachi. 2006. Measuring Gaze Direction Perception Capability of Humans to Design Human Centered Communication Systems. *Presence: Teleoperators and Virtual Environments* 15, 2 (2006), 123–138. <https://doi.org/10.1162/pres.2006.15.2.123>
- [27] Rob Jacob and Sophie Stellmach. 2016. What You Look at is What You Get: Gaze-Based User Interfaces. *Interactions* 23, 5 (Aug. 2016), 62–65. <https://doi.org/10.1145/2978577>
- [28] David Kirk, Tom Rodden, and Danaë Stanton Fraser. 2007. Turn It This Way: Grounding Collaborative Action with Remote Gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '07*). ACM, New York, NY, USA, 1039–1048. <https://doi.org/10.1145/1240624.1240782>
- [29] David S. Kirk and Danaë Stanton Fraser. 2005. The Effects of Remote Gesturing on Distance Instruction. In *Proceedings of Th 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The next 10 Years!* (Taipei, Taiwan) (*CSCL '05*). International Society of the Learning Sciences, Bloomington, Indiana, USA, 301–310.
- [30] Takeshi Kurata, Nobuchika Sakata, Masakatsu Kouroggi, Hideaki Kuzuoka, and Mark Billinghurst. 2004. Remote collaboration using a shoulder-worn active camera/laser. In *Eighth International Symposium on Wearable Computers (ISWC '04)*. IEEE, Washington, DC, USA, 62–69.
- [31] Ka-Wai Kwok, Loi-Wah Sun, George Mylonas, David James, Felipe Orihuela-Espina, and Guang-Zhong Yang. 2012. Collaborative Gaze Channelling for Improved Cooperation During Robotic Assisted Surgery. *Annals of biomedical engineering* 40 (05 2012), 2156–67. <https://doi.org/10.1007/s10439-012-0578-4>
- [32] Morgan Le Chénéchal, Thierry Duval, Valérie Gouranton, Jérôme Royan, and Bruno Arnaldi. 2019. Help! I Need a Remote Guide in My Mixed Reality Collaborative Environment. *Frontiers in Robotics and AI* 6, Article 106 (2019), 16 pages. <https://doi.org/10.3389/frobt.2019.00106>
- [33] Jerry Li, Mia Manavalan, Sarah D'Angelo, and Darren Gergle. 2016. Designing Shared Gaze Awareness for Remote Collaboration. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (San Francisco, California, USA) (*CSCW '16 Companion*). ACM, New York, NY, USA, 325–328. <https://doi.org/10.1145/2818052.2869097>
- [34] Lars Lischke, Valentin Schwind, Robin Schweigert, Paweł W. Woundefinedniak, and Niels Henze. 2019. Understanding Pointing for Workspace Tasks on Large High-Resolution Displays. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia* (Pisa, Italy) (*MUM '19*). ACM, New York, NY, USA, Article 25, 9 pages. <https://doi.org/10.1145/3365610.3365636>
- [35] Romy Müller, Jens Helmert, and Sebastian Pannasch. 2014. Limitations of gaze transfer: Without visual context, eye movements do not help to coordinate joint action, whereas mouse movements do. *Acta Psychologica* 152 (10 2014), 19–28. <https://doi.org/10.1016/j.actpsy.2014.07.013>
- [36] Romy Müller, Jens Helmert, Sebastian Pannasch, and Boris Velichkovsky. 2013. Gaze transfer in remote cooperation: Is it always helpful to see what your partner is attending to? *Quarterly journal of experimental psychology* (2006) 66, 7 (07 2013), 1302–1316. <https://doi.org/10.1080/17470218.2012.737813>
- [37] Norman Murray, David Roberts, Anthony Steed, P.M. Sharkey, Peter Dickenson, John Rae, and Robin Wolff. 2009. Eye Gaze in Virtual Environments: Evaluating the need and initial work on implementation. *Concurrency and Computation Practice and Experience* 21, 11 (01 2009), 1437–1449.
- [38] Joshua Newn, Singh Ronal, Fraser Allison, Madumal Prashan, Eduardo Velloso, and Frank Vetere. 2019. Designing Interactions with Intention-Aware Gaze-Enabled Artificial Agents. *Human-Computer Interaction – INTERACT 2019* 11747, 0 (08 2019), 255–281. https://doi.org/10.1007/978-3-030-29384-0_17

- [39] Joshua Newn, Eduardo Velloso, Fraser Allison, Yomna Abdelrahman, and Frank Vetere. 2017. Evaluating Real-Time Gaze Representations to Infer Intentions in Competitive Turn-Based Strategy Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) (*CHI PLAY '17*). ACM, New York, NY, USA, 541–552. <https://doi.org/10.1145/3116595.3116624>
- [40] Mai Otsuki, Keita Maruyama, Hideaki Kuzuoka, and Yusuke SUZUKI. 2018. Effects of Enhanced Gaze Presentation on Gaze Leading in Remote Collaborative Physical Tasks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). ACM, New York, NY, USA, 1–11. <https://doi.org/10.1145/3173574.3173942>
- [41] Jiazhi Ou, Xilin Chen, Susan R. Fussell, and Jie Yang. 2003. DOVE: drawing over video environment. In *Proceedings of the Eleventh ACM International Conference on Multimedia*. ACM, New York, NY, USA, 100–101.
- [42] Jiazhi Ou, Lui Min Oh, Jie Yang, and Susan R. Fussell. 2005. Effects of task properties, partner actions, and message content on eye gaze patterns in a collaborative task. In *Proceedings of the 2005 CHI Conference (CHI '05)*. ACM, New York, NY, USA, 231–240.
- [43] Jiazhi Ou, Lui Oh, Susan R. Fussell, Tal Blum, and Jie Yang. 2008. Predicting Visual Focus of Attention From Intention in Remote Collaborative Tasks. *IEEE Transactions on Multimedia* 10 (11 2008), 1034–1045. Issue 6. <https://doi.org/10.1109/TMM.2008.2001363>
- [44] Jiazhi Ou, Lui Min Oh, Susan R. Fussell, Tal Blum, and Jie Yang. 2005. Analyzing and Predicting Focus of Attention in Remote Collaborative Tasks. In *Proceedings of the 7th International Conference on Multimodal Interfaces* (Toronto, Italy) (*ICMI '05*). ACM, New York, NY, USA, 116–123. <https://doi.org/10.1145/1088463.1088485>
- [45] Doug Palmer, Matt Adcock, Jocelyn Smith, Matthew Hutchins, Chris Gunn, Duncan Stevenson, and Ken Taylor. 2007. Annotating with Light for Remote Guidance. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces* (Adelaide, Australia) (*OZCHI '07*). ACM, New York, NY, USA, 103–110. <https://doi.org/10.1145/1324892.1324911>
- [46] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Barrett Ens, and Mark Billinghurst. 2017. [POSTER] CoVAR: Mixed-Platform Remote Collaborative Augmented and Virtual Realities System with Shared Collaboration Cues. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. IEEE, USA, 218–219. <https://doi.org/10.1109/ISMAR-Adjunct.2017.72>
- [47] Thammathip Piumsomboon, Lee Gun A., Hart Jonathon D., Barrett Ens, Lindeman Robert W., Thomas Bruce H., and Mark Billinghurst. 2018. Mini-me: an adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, Anna Cox and Mark Perry (Eds.). ACM, New York, NY, USA, Article 46, 13 pages. <https://doi.org/10.1145/3173574.3173620> International Conference on Human Factors in Computing Systems 2018, CHI 2018 ; Conference date: 21-04-2018 Through 26-04-2018.
- [48] John Rae, William Steptoe, and David Roberts. 2011. Some Implications of Eye Gaze Behavior and Perception for the Design of Immersive Telecommunication Systems. In *Proceedings - IEEE International Symposium on Distributed Simulation and Real-Time Applications*. IEEE, USA, 108–114. <https://doi.org/10.1109/DS-RT.2011.37>.
- [49] Nobuchika Sakata, Nobuchika Kurata, Takekazu Kato, Masakatsu Kourog, and Hideaki Kuzuoka. 2003. WACL: supporting telecommunications using - wearable active camera with laser pointer. In *Proceedings of the Seventh IEEE International Symposium on Wearable Computers*. IEEE, USA, 53–56.
- [50] William Steptoe, Oyewole Oyekoya, Alessio Murgia, Robin Wolff, John Rae, Estefania Guimaraes, David Roberts, and Anthony Steed. 2009. Eye Tracking for Avatar Eye Gaze Control During Object-Focused Multiparty Interaction in Immersive Collaborative Virtual Environments. In *Proceedings - IEEE Virtual Reality*. IEEE, USA, 83–90. <https://doi.org/10.1109/VR.2009.4811003>
- [51] William Steptoe, Robin Wolff, Alessio Murgia, Estefania Guimaraes, John Rae, Paul Sharkey, David Roberts, and Anthony Steed. 2008. Eye-Tracking for Avatar Eye-Gaze and Interactional Analysis in Immersive Collaborative Virtual Environments. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (San Diego, CA, USA) (*CSCW '08*). ACM, New York, NY, USA, 197–200. <https://doi.org/10.1145/1460563.1460593>
- [52] Matthew Tait and Mark Billinghurst. 2015. The Effect of View Independence in a Collaborative AR System. *Computer Supported Cooperative Work (CSCW)* 24 (2015), 563–589. <https://doi.org/10.1007/s10606-015-9231-8>
- [53] Vincent van Rheden, Bernhard Maurer, Dorothé Smit, Martin Murer, and Manfred Tscheligi. 2017. LaserViz: Shared Gaze in the Co-Located Physical World. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction* (Yokohama, Japan) (*TEI '17*). ACM, New York, NY, USA, 191–196. <https://doi.org/10.1145/3024969.3025010>
- [54] Peng Wang, Xiaoliang Bai, Mark Billinghurst, Shusheng Zhang, Weiping He, Dechuan Han, Yue Wang, Haitao Min, Weiqi Lan, and Shu Han. 2020. Using a Head Pointer or Eye Gaze: The Effect of Gaze on Spatial AR Remote Collaboration for Physical Tasks. *Interacting with Computers* 32 (07 2020), 153–169. Issue 2. <https://doi.org/10.1093/iwcomp/iwaa012>
- [55] Peng Wang, Shusheng Zhang, Xiaoliang Bai, Mark Billinghurst, Weiping He, Shuxia Wang, Xiaokun Zhang, Jiaxiang Du, and Yongxing Chen. 2019. Head Pointer or Eye Gaze: Which Helps More in MR Remote Collaboration?. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, USA, Article 8798024, 2 pages. <https://doi.org/10.1109/VR.2019.8798024>
- [56] Aiden Wickey and Leila Alem. 2007. Analysis of Hand Gestures in Remote Collaboration: Some Design Recommendations. In *Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces* (Adelaide, Australia) (*OZCHI '07*). ACM, New York, NY, USA, 87–93. <https://doi.org/10.1145/1324892.1324909>
- [57] Takuya Yamamoto, Mai Otsuki, Hideaki Kuzuoka, and Yusuke Suzuki. 2018. Tele-Guidance System to Support Anticipation during Communication. *Special issue "Spatial Augmented Reality" of Multimodal Technologies and Interact* 3, Article 55 (2 2018), 12 pages. <https://doi.org/10.3390/mti2030055>
- [58] Yanxia Zhang, Ken Pfeuffer, Ming Ki Chong, Jason Alexander, Andreas Bulling, and Hans-Werner Gellersen. 2016. Look together: using gaze for assisting co-located collaborative search. *Personal and Ubiquitous Computing* 21 (2016), 173–186. <https://doi.org/10.1007/s00779-016-0969-x>