

Effects of Limited Field of View on Musical Collaboration Experience with Avatars in Extended Reality

Suibi Che-Chuan Weng
ATLAS Institute
University of Colorado, Boulder
Boulder, USA
suibi.weng@colorado.edu

Torin Hopkins
SolJAMM Research
Denver, USA
0000-0001-7359-2906

Shih-Yu Ma
Department of Computer Science
University of Colorado Boulder
Boulder, USA
shih-yu.ma@colorado.edu

Amy Banić
Interactive Realities Lab
Department of Electrical Engineering and Computer Science
University of Wyoming
Laramie, USA
abanic@uwyo.edu

Ellen Yi-Luen Do
ATLAS Institute
University of Colorado, Boulder
Boulder, USA
ellen.do@colorado.edu

Abstract—During musical collaboration, visual cues are essential for communication between musicians. Extended Reality (XR) applications, often used with head-mounted displays like Augmented Reality (AR) glasses, can limit the field of view (FOV) of players. We conducted a study to investigate the effects of limited FOV on co-presence, gesture recognition, overall enjoyment, and reaction time.

Initially, we observed experienced musicians collaborating informally with and without visual occlusion, noting that collaboration suffered with limited FOV. We then conducted a within-subjects study with 19 participants, comparing an unrestricted FOV holographic setup called HoloJam to Nreal AR glasses with a 52° limited FOV. In the AR setup, we tested two conditions: standard AR with a 52° FOV and a modified AR notification system called Mini Musicians.

Results showed that HoloJam provided higher co-presence, quicker gesture recognition, and greater enjoyment. The Mini Musicians application reduced reaction time and maintained enjoyment compared to the standard AR setup. We conclude that limited FOV impacts musical collaboration, but notifications can improve reaction time and should be considered in future XR music collaborations.

Index Terms—Musical Metaverse, Musical XR, Digital Audio Workstation, encumbered interactions

I. INTRODUCTION

During the COVID-19 pandemic, individuals continued to seek ways to connect with others through music, including engaging in networked music platforms [1]–[3]. Remote music collaboration software such as Jacktrip [3] and Jamulus [2] facilitated playing music together at a distance and aimed to recreate an authentic auditory musical experience. While networked video facilitates visual collaboration [3], the use of Extended Reality (XR) applications presents an alternative avenue to explore social dynamics in three dimensions [4], [5].

In musical collaboration, communication between players occurs spatially, usually involving directed hand gestures, facial expressions, and eye-gaze [1]. This information is essential for musical transitions and collective musical coordination.

Recently, the use of AR headsets has seen a substantial increase, sparking a heightened interest in the study of interactions between humans and virtual content in augmented reality. As AR technology progresses, understanding how people perceive and interact with virtual elements becomes increasingly important, despite current limitations like a restricted field of view (FOV). Research in this area often examines how users process information that lies outside of their FOV [6], [7]. Much of this research focuses on enhancing task performance through message notifications rather than improving the social presence of avatars outside the FOV [8]–[11]. Musical collaboration, which heavily depends on communicative gestures, provides a valuable context for studying interactions with avatars [1], [12]. While gestures like hand movements and eye contact are easily observed in face-to-face interactions, detecting these social cues becomes challenging in multi-avatar collaborations when avatars are outside the user's FOV.

In our study, we provided a more detailed investigation building upon the research by Weng et al. [13], [14] regarding the impact of FOV limitations on musical collaboration.

To delve deeper into this problem, we carried out an investigation in three parts, shown in Figure 2. Insights from these sessions and previous work on AR notifications led researchers to test a holographic musician system against a known type of peripheral view notification—blue spheres. [9], [13]. Researcher noticed that reaction times were shorter when participants used the notification system, although they reported feeling distracted. [13]

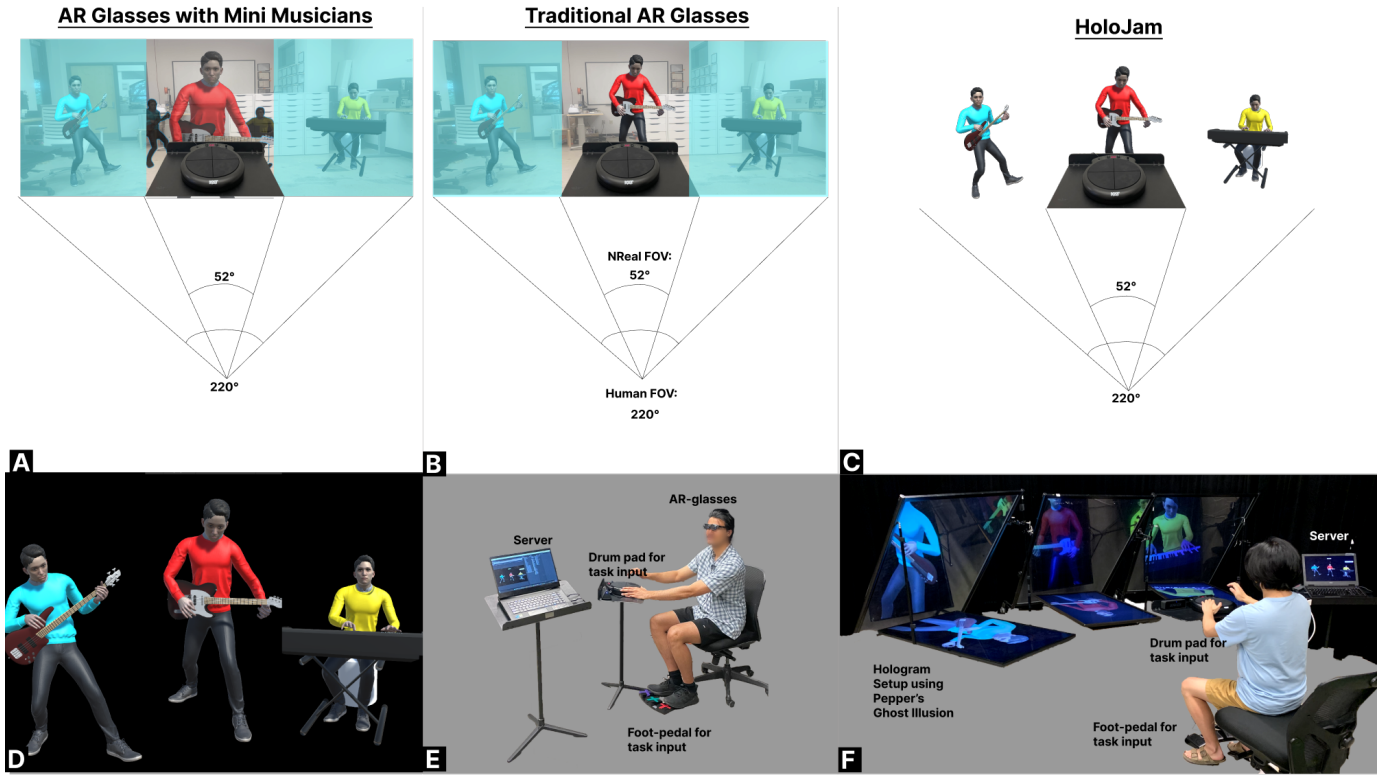


Fig. 1. A) A representation of the experimental setup using AR glasses, a MIDI drum, and mini musicians; B) An illustration of limited field of view with no mini musicians; C) Demonstrates the holographic setup with no limit on field of view.

Additionally, we observed live jam sessions where musicians played music together under two conditions: 1) normally, and 2) while wearing glasses that blocked their peripheral vision. Building on our observations and previous related work, we developed three systems to test our hypothesis. The first system features HoloJam which was a three holograms created using the Pepper's ghost illusion, representing a condition without any FOV limitations (Figure 1 C). The second condition uses the exact same application but uses AR glasses (Figure 1 A and B) and user's position, with a limited FOV of 52°. The third and final condition is the same application, but when avatars are outside the field of view a small and semi-transparent version of the avatar appears within the field of view (mini musicians). The notification system was called Mini Musicians. We conducted a within-subject study comparing the three conditions. In the study participants were asked to play a hand drum with three avatar musicians and indicate when the avatar is looking at them.

Throughout this paper we discuss work that inspired our experimental design, methods and results, and implications for future work in the field. Results showed that a limited FOV does affect musical collaboration experience. Conditions associated with a larger field of view yielded a significantly higher score in sense of co-presence with the avatars and overall enjoyment. However, the notification strategy of mini musicians reduced the reaction time significantly with AR glasses conditions.

Contributions of our study are:

- A limited FOV significantly reduces musicians ability to communicate during a jam session, which impacts the music and perceived experience
- The mini musicians notification helps to reduce the reaction time with AR glasses
- Notifications while playing music in AR do not significantly affect co-presence, likely because of distraction

II. RELATED WORK

Our research focuses on three distinct areas. First, we delve into the realm of musical collaboration in extended reality (XR), examining how avatars can enhance the immersive experience for musicians engaged in cooperative musical activities. Second, we explore the development of notifications specifically designed for AR glasses, aiming to address the challenge of objects that are outside the user's field of view. By providing visual cues or indicators for these objects, we aim to enhance the overall AR experience and improve user engagement. Lastly, we investigate the concept of social presence in avatars within the AR context, examining how users perceive and interact with avatars in virtual environments, with a particular emphasis on their impact on cooperative interactions and perceived presence.

A. Musical collaborations in XR

Over the past two decades, the intersection of music and XR has become a well-established area of research, transforming

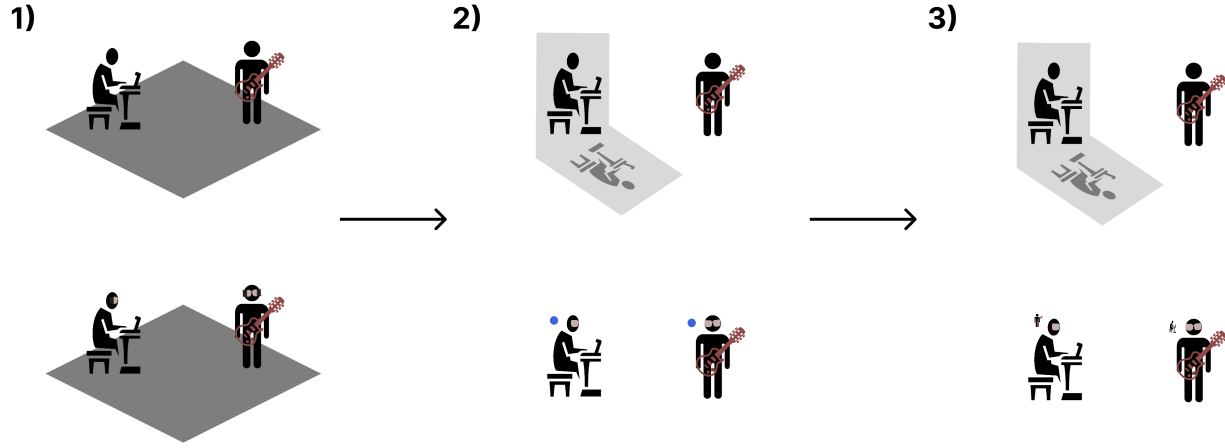


Fig. 2. (1) Assessing in person jam sessions with and without visual occlusion at 52 degrees (2) Assessment of holograms vs blue light notification, [13] (3) Assessment of holograms vs mini musicians [14]

traditional musical interactions by enabling players to engage with virtual objects, agents, and environments [15]. During the COVID-19 pandemic, face-to-face music collaboration was limited, leading to a burgeoning exploration of mixed reality technologies as a means to create a sense of presence and community. The project *Wish You Were Here* [16], is a laboratory study conducted to assess the potential of mixed reality for music collaboration, comparing it to an audio-only baseline. The results showed increased co-presence and positive affect for musicians jamming in mixed reality.

Various other methods have been explored for collaborative music-making, such as using a composer table [17], [18] or employing a graphic interface to document the collaborative experience [19], [20]. However, the use of AR or virtual reality (VR) has demonstrated superior immersive experiences in musical collaborations [12], [21]. Studies, including the AR drum circle [1], have emphasized the significance of social cues in XR musical collaborations, where animated avatars enhance the immersive experience. Nevertheless, challenges remain, such as the limited FOV of AR glasses, which can lead to virtual characters being cut off, affecting communication and musical performance. Our study aims to provide statistical insights into the impact of restricted FOV on musical collaborations with avatars via AR glasses.

B. Notifications for limited FOV of AR glasses

Many optical-see-through displays have a relatively narrow field of view. However, a limited FOV can constrain how information can be presented and searched through [6]. Researchers have proposed solutions at both the hardware and software level to address the challenges of limited field of view in AR applications.

Hardware solutions aim to expand the user's peripheral vision. For example, *RadialLight* [9] utilizes LED plugins to

extend the field of view, while projects like *Eye-q* [22] introduce wearable peripheral displays embedded in eyeglasses.

On the software side, visual notifications have been employed to mitigate the limitations of limited field of view. Projects such as the *Flying Arrow* [8] utilize a moving arrow indicator to direct the user's attention to out-of-screen objects. Matsuzoe et al. [23] explored the use of vibrating icons to indicate the proximity of out-of-screen targets. Advancing this research further, Qian et al. [10] tested different notification approaches in search tasks using AR glasses and found that edge notifications with a blue indicator yielded better performance.

In addition to notifications, projects like *Push the Red Button* [24] have shown the effectiveness of wrist-based notifications when combined with other AR elements. Notifications displayed within the user's field of view are quickly noticed and understood if the user's view direction is known. Meanwhile, the introduction of a mobile peripheral vision model, as studied in *The Peripheral Vision* study [25], aims to simplify multitasking on smart glasses without requiring additional hardware. Rzaev et al.'s research [26] focuses on displaying notifications during face-to-face communication, which further highlights the importance of tailoring notification strategies for different scenarios within AR applications.

While previous research has primarily focused on navigation tasks and object finding, our study specifically targets how users perceive avatar behavior in musical activities. Drawing from previous notification research we measure reaction time, co-presence [7], [27], and enjoyment as measurements of effective notification strategies for musical tasks in XR. Leveraging software notifications with AR glasses, we aim to address the limitations inherent to AR glasses and enhance the musical collaboration experience.



Fig. 3. (A) limited FOV Goggles (B) Normal Jam Session: Musicians with full FOV (C) limited FOV: Musicians Jamming with the Goggles (D) Jam Session

C. Social presence with the virtual Avatars in AR

The importance of interacting with avatars in AR research is highlighted, and understanding users' perception of avatar presence is crucial for assessing its impact in musical collaboration [28]. Previous studies have utilized social presence questionnaires to evaluate the sense of presence [1], [7], [27], [29]. Osmers et al. [30] conducted a review study supporting the influence of social presence in asymmetric cooperation settings on user preferences and acceptance.

Furthermore, Hart et al. [31] found a preference for 3D avatars in XR communication scenarios, leading us to focus on 3D avatar musicians in our study. In an AR drum circle, participants paid a lot of attention to the avatar's body expression, as observed by Hopkins et al. [1]. These findings inspired us to explore software modifications for AR such that the avatar's full body appears on the external screen.

Similarly, eye contact and body gestures have been shown to improve task performance in several studies [32]–[36]. However, the limited field of view (FoV) in AR diminishes avatar social presence and the user's perception of the avatar's message [37]. Shin et al. [37] found that users wearing AR head-mounted displays (HMDs) faced challenges due to frequent head movements caused by limited FOV. Piumsomboon et al.'s Mini Me project [7], which compared social presence with AR glasses between a human-size avatar and a mini avatar on a screen in various scenarios, showed that Mini Me achieved a higher social presence score. Their research scenarios included urban planning, design tasks [33], meeting rooms [32], and tea parties [7].

In our research, we specifically focus on musical activities with multiple avatars that are outside of the FOV. Building upon *prior research, observations of a musical jam session, and player feedback*, we created a **Mini Musicians Avatar** that appears when the life-sized avatar is outside of the player's FOV.

III. METHODS

In our study we investigate how FOV affects musical collaboration. First, we observe a live music jamming session with and without glasses that restrict the FOV to 52°. We recorded the whole session and analyzed the video. Post-analysis and interview, we built the experiment apparatus according to the analyzed result from the jam session. This method of observation was inspired by musician video observation methods conducted by Hopkins, et al [1].

A. Observation of a Jamming Session

To examine the impact of limited field of view (FOV), we conducted two 30-minute jam sessions with two guitar players, a bass player, a drummer, and a singer (Figure 3 D). The first session was a control with no external factors (Figure 3 B). In the second session, participants wore goggles restricting their vision to 52°, simulating AR goggles (Figure 3 C). Both sessions were recorded with three web cameras from different angles. We assessed music quality, performance coherence, and logged participants' body positioning and language. Two minutes of each session were analyzed per participant by a research assistant using Datavyu and written notes [38].

1) *Musicians*: The jam session group consisted of five musicians aged 25 to 31: two guitarists, a bassist, a keyboardist, and a drummer. Each had over five years of experience with their instruments and regularly jammed together once or twice a week.

2) *Normal jamming*: During a regular jamming session, the musical coordination was excellent, with a high level of synchronization and a rapid tempo. Although there were a few minor miscommunications, they were resolved swiftly. The participants engaged in frequent communication, often making quick eye contact and occasionally exchanging minimal verbal cues. Throughout the session, each musician maintained multiple instances of eye contact with other musicians and their respective instruments.

3) *Limited FOV jamming*: In the jam session with a limited FOV, the musical collaboration was observably worse compared to the previous session. The music played had reduced coherence and experimentation, as all the musicians adopted a more conservative approach. Communication during this session was also hindered, with fewer instances of actual eye contact among the musicians. In fact, some musicians stopped looking up altogether. Additionally, there was an increase in overall body movement, including frequent head turning. For instance, the drummer had to move their head extensively to see the drums properly, affecting their playing.

4) *Summary of Observation*: Between the two, musical coherence was better in the control test over the test with the vision-restricting goggles on. The body language was also more open (head up, bodies turned toward each other, etc) in the first test. Participants looked down a lot more with the goggles on and made less eye contact. It was also observed that the drummer struggled to see their entire drum kit with the goggles on as the kit was wider than their FOV.

The number of times participants either looked up or around at other participants were further analyzed. This was defined as clearly looking over to another participant's instrument or eyes. For this test, 2 minutes of each session were watched, observing one participant. Then, it was repeated for the same 2 minutes of the session for the other participants. For the control session, the participants had a total of **62** times when they looked at another participant. For the other session, the participants had a total of **27** times when they looked at another participant. This shows that the limited FOV caused less non-verbal communication, due to a lack of peripheral vision.

5) *Summary of Post-Jam Session Interview:* We conducted an analysis of interviews with the participants regarding their experiences of jamming with limited FOV compared to normal jam sessions. The main takeaways from the interviews are as follows:

Control Jam: Participants who experienced a normal jamming session reported that the jam went well and felt sufficiently representative of their usual jams.

Limited FOV Jam: Initially, participants who jammed with limited FOV did not feel much difference compared to the control jam. However, they gradually started noticing certain limitations. Specifically, they mentioned that they found themselves focusing straight ahead and not looking around. This narrowed focus caused some communication breakdowns as players would often take longer to notice that someone was looking at them or trying to convey something.

Missing Visual Cues: One participant expressed frustration, stating that they don't know what they are missing because they cannot see it due to the limited FOV. This inability to perceive the entire visual scene hindered their overall experience and understanding of the jamming session.

Forgetting Presence: Another noteworthy observation was that participants mentioned how, with the limited FOV, they sometimes forgot that certain players were present in the virtual space. This forgetfulness could potentially lead to unintentional exclusion or overlooking the contributions of certain individuals during the jam.

Desire for Mini Avatars: Three participants mentioned that they would love to see a mini version of the avatars, similar to those in Mario Kart. They suggested that having smaller avatars displayed in a corner of their field of view could help them maintain awareness of other players and enhance communication during the jam.

Overall, the interviews revealed that while the limited FOV initially did not feel drastically different, participants gradually became aware of the constraints it imposed. The reduced ability to observe the entire scene and the potential for communication breakdowns were notable drawbacks. Participants expressed a desire for alternative visual cues, such as mini avatars, to enhance their awareness and communication in the limited FOV setting. These insights provide valuable feedback for the design and improvement of collaborative AR experiences, particularly in the context of musical activities. We decided to build and test a mini avatar system, as it

was consistent with previous work [7] and desired by several interviewed musicians.

B. Limitations of Prior Work

From previous work, Weng et al. [14] conducted an experiment testing the limitations of FOV. Their results indicated that the holographic display without FOV limitations achieved the highest co-presence score, the lowest response time, and the highest task enjoyment score. This suggests that FOV limitations negatively affect musical collaboration with avatars. Additionally, while AR glasses with notifications may improve co-presence, they could potentially reduce enjoyment. In subsequent work, Weng et al. [13] enhanced the notifications by introducing "mini musicians." In our paper, we present more details on the system implementation and additional results from our user study.

IV. RESEARCH QUESTIONS

The key insights from prior studies are summarized below: **XR technology has transformed traditional musical interactions**, allowing musicians and audiences to engage with virtual objects and environments. Mixed reality technologies were explored as a means to create a sense of presence for musicians.

Social presence with virtual avatars in AR is crucial for assessing their impact on musical collaborations. Previous research has shown the significance of social presence in asymmetric cooperation settings, and avatars in 3D form have been preferred in XR communication scenarios.

Opportunity to study musical activities with multiple avatars, examining the impact of avatar behavior on user perception and exploring software notifications with AR glasses to address their limitations.

Building upon these findings, our research questions are:

- RQ1: How fast can a player notice an avatar musician when the avatar is communicating with the player non-verbally (avatar is looking up at the player) with and without a limited FOV?
- RQ2: What are the Aggregated Social Presence Scores (given by Aggregated Social Presence Scores questionnaire [27]) with the avatars with and without a limited FOV? [13]
- RQ3: Does a limited FOV affect task enjoyment?

V. STUDYING THE EFFECTS OF NOTIFICATIONS IN LIMITED FOV MUSICAL SCENARIOS

To answer our research questions, we redesigned the previous pilot study and conducted a user study using our improved system.

A. System design: HoloJam and Mini Musicians

According to the interview and previous research we built two XR hardware setups enabling players to jam with virtual avatars. As successfully implemented in pilot studies, the first system was created without a limited FOV using holograms. The second system used AR glasses running an application

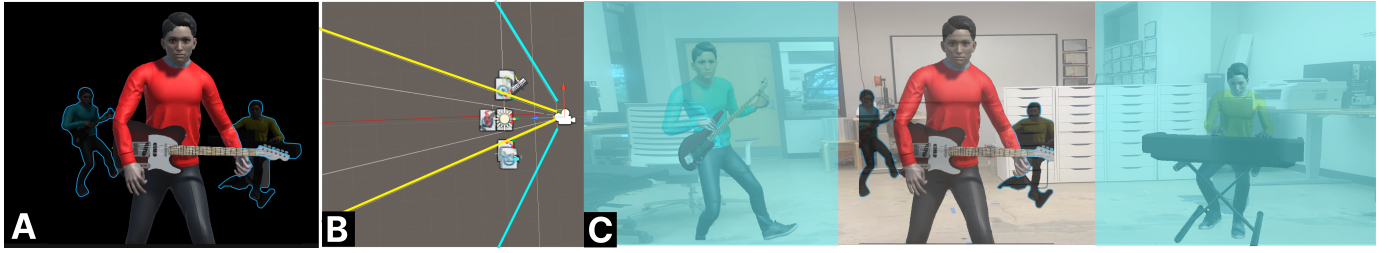


Fig. 4. (A) Virtual Camera views (B) Virtual Cameras setup Yellow lines: The range of the limited FOV Blue lines: A wider range of FOV (C) Mini musicians: AR glasses view

with a limited FOV of 52° . Using AR glasses we had two conditions: one with mini musicians (Figure 1 A) and one without any additional content (Figure 1 B).

B. HoloJam

The first setup was a Hologram Jam (Figure 1 C), utilizing a Pepper's ghost illusion. We chose this setup over a screen display because previous research indicates that the Pepper's ghost illusion provides a more immersive experience, particularly in a musical performance environment [39]–[41]. Furthermore, according to the works of XR Jam [4] and Schlagowski et al. [16], musicians tend to move around while playing. The hologram setup offers two important features: it gives users a sense of real space and allows in-person players to perform alongside avatars without being occluded by a screen. (e.g. stands behind the avatars) The hardware setup consisted of a laptop PC equipped with an Intel 10700k CPU, an Nvidia 1070 GPU, and 32GB of memory. We used a transparent acrylic board placed at a 45-degree angle from the screen to reflect the television's visuals as a hologram screen. This setup enabled participants to interact with the avatars without the limitations of a restricted FOV.

C. AR Glasses

The second was a jam session using NReal AR glasses (Figure 1 E) ¹. In the AR glasses set up, we placed three virtual avatars in the AR environment. One in the middle, and the other two avatars on both sides beyond 52° (outside FOV).

1) *Notification Design Mini Musicians*: The notification method we used was the Mini Musicians. This design was adapted from the Mini Me research [7], which we modified for a musical context to enhance the user experience in musical collaborations.

In our virtual environment setup, we implemented a second virtual camera with a wider FOV than the main camera (Figure 4 B). This second camera was placed on a 2D transparent canvas positioned in front of the main camera. The main camera displayed the primary view (Figure 4 A), while the second camera provided an extended view, showing smaller objects (in this case, the avatars) that were outside the FOV of the main camera as the user moved their head to look around (Figure 4 C).

This camera setup offers for AR developers. By incorporating a second virtual camera, developers can adapt the off screen notification system to various contexts and content without needing to redesign the entire notification method. This approach ensures that users maintain a wider vision in the AR view, enhancing their ability to interact with and respond to virtual elements during musical collaboration. Additionally, this method provides a more immersive experience by maintaining awareness of avatars outside the immediate FOV, thereby improving the overall effectiveness of AR notifications.

D. Interaction and Content Design

1) *Interactive Device for the Participants*: In this apparatus, we provided two input devices for the participants. The first one was a MIDI drum pad that played the drum sound from a speaker. The second one was a foot pedal (Figures 1 E and F) that indicated three positions of the avatar which let the user select by step on the cross response position of the pedal, which avatar was following them.

2) *XR Content design*: We designed the musical collaboration scenarios that enabled participants to play musical instruments with virtual avatars. It was a jam session with three animated avatars playing different instruments: bass, guitar and Keyboard (Figure 1 D). The animations reacted to the background music. In both scenarios the avatar models and several animations were downloaded from Mixamo ². The reactive animations were controlled by an animated rig plugin in the Unity Game Engine ³.

VI. USER EVALUATION

According to the pilot study results we improved our AR notification system by solely including mini musicians, rather than light or symbolological indicators. Furthermore, we simplified the experiment procedure to reflect a more thorough design approach and mechanism of data collection.

A. Experimental Design

To address our research questions we conducted an experiment that compared musical experience and social presence with avatars in different FOV situations in AR. We conducted a 2x3 within-subjects experiment to simulate a jam session with

¹<https://www.nreal.ai/light/>

²<https://www.mixamo.com/>

³<https://unity.com/>

three avatars. In the AR conditions, we used two contexts: 1) visual notifications use mini musicians to indicate avatar gesture when outside of the FOV. The avatars have a single gesture, to look up at the participant from the default position of looking at the avatar's own instrument, and 2) no visual notification. In the holographic setup, we put three holograms in front of a participant with the same FOV as the AR glasses conditions. Like previous experiments the hologram condition simulates playing the Jam session with avatars, but without an FOV limitation. We let participants play a midi drum because the drum playing is easier to a regular participant than other melody instruments.

B. Participant

A total of 19 participants were recruited for the study through email lists and word of mouth. All participants were university students, and the experiment took place in a research laboratory on the university campus. Each participant got 5 dollars cash compensation for their time. Among the participants, 11 were undergoing professional music training. The number of training years varied widely, ranging from 0 to over 10 years. The majority of participants had 2-3 years of training.

C. Procedure

The experiment ensured informed consent and explained the drumming task, equipment, and data handling. Participants synchronized drumming with music while identifying an avatar looking at them (pressing a pedal button).

Three conditions were tested: hologram (participants adjusted their position in front of 3 holograms with lights off), and two AR glasses setups (researchers helped participants wear glasses and adjusted the environment).

A counterbalanced Latin square design ensured each participant experienced all conditions in a random order (each lasting 8 minutes). Questionnaires assessed social presence, enjoyment, and post-test feedback. The entire session averaged 45-50 minutes.

VII. RESULTS

We measured three major quantitative variables. The first was reaction time, measured as the amount of time it took the participants to notice an avatar was looking at them. The second was enjoyment, which employed a 1-7 likert scale to measure whether the participants were enjoying the tasks or not. The last quantitative measure obtained was the social presence aggregate score. Finally, we collected the qualitative data from the participants' free writing of at least 150 words [1], [42].

A. Reaction Time

To answer our *RQ1*, we measured the response time when the participants noticed the following avatar. We collected 10 rounds of data from each participant and measured the mean response time in milliseconds after all rounds. To analyze the data, a repeated measures Analysis of Variance (RM-ANOVA) was conducted on the data. The result showed that

mean enjoyment differed significantly across three conditions $F(2, 36) = 4.476, p = 0.018$ (Figure 5 B). The hologram condition had significantly faster reaction times than the AR glasses condition without notifications ($\eta_p^2=0.34$). As well as AR glasses with mini musicians was also significantly faster than the no notification condition ($\eta_p^2=0.20$). Based on the results, we can say that the limited FOV does affect the time it takes for the user to notice the avatar (**RQ1**). Additionally, the results show that the Mini Musicians helped participants notice the avatar faster than in the no-notification condition.

B. Social Presence Aggregate Score

In the previous works Weng et al. [13] address *RQ2*, researchers assessed the Social Presence Aggregate Score, finding significant differences in all categories (Co-Presence, PMU, PBI) across conditions, particularly favoring the hologram condition. The hologram condition consistently scored higher, indicating that a limited FOV impacts the social presence in a musical collaboration. No significant differences were noted between Mini Musicians and No Notifications.

C. Enjoyment

After the participants finished each task, we gave a task enjoyment questionnaire. Quantitative scores between 1 and 7 (1 - lowest, 4 - neutral, and 7 - highest) were indicated for measure enjoyment. The result showed that mean reaction time differed significantly across three conditions $F(2, 36) = 13.698, p \leq 0.001$ (Figure 5 A). The hologram condition had significantly better enjoyment scores than the AR glasses condition without notifications ($\eta_p^2 \leq 0.01$) and the AR glasses condition with mini musicians ($\eta_p^2 \leq 0.01$). There were no significant differences between Mini Musicians and Without notifications. In this result, users without a limited field of view (FOV) experience better enjoyment with the jamming experience (**RQ3**).

D. Participant responses

The analysis of participants' free writing response from the three jamming experiences—AR jamming, AR jamming with Mini Musicians notifications, and HoloJam—revealed significant differences in terms of field of view, immersion, comfort, and preference. The AR jamming experience had limitations related to the narrow field of view and discomfort caused by wearing glasses. Despite these limitations, users still found the AR experience immersive and appreciated the detailed rendering of avatars. The introduction of notifications in AR jamming aimed to address the field of view issue but received mixed responses. While some users found the notifications helpful, others found them distracting or ineffective. In contrast, HoloJam offered the most natural and immersive experience, with a broader FOV, better avatar interaction, and increased comfort. Users preferred the HoloJam experience over the AR versions due to its convenience and realism.

The implementation of notifications in the AR application aimed to improve the limited field of view and bridge the gap with the holographic experience. However, the feedback

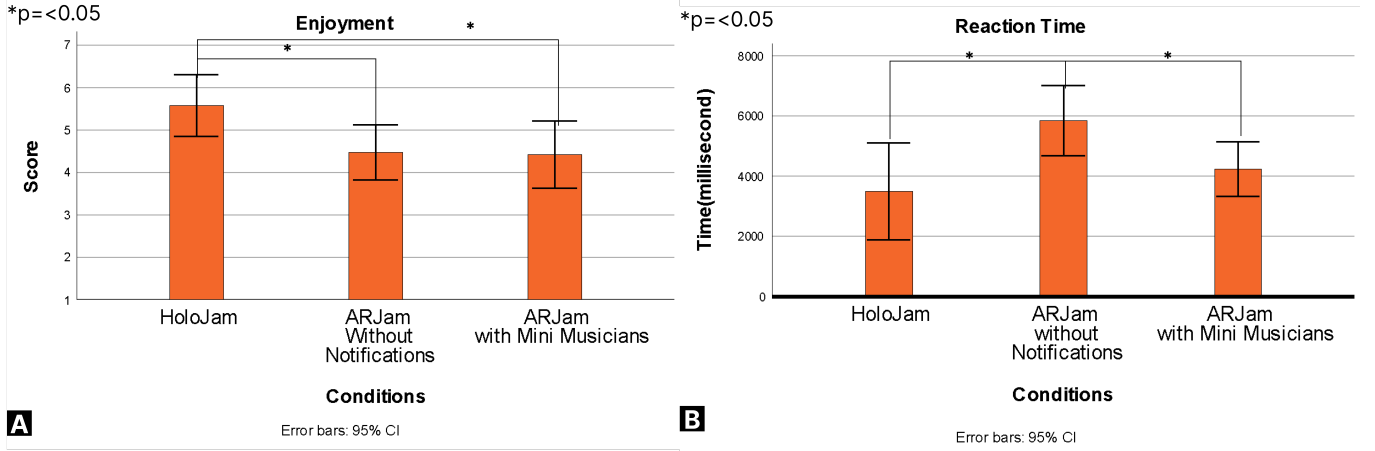


Fig. 5. (A)Enjoyment:Enjoyment score (B)Reaction Time: the time a participant took to notice an avatar was attempting to communicate

indicated that the notifications did not significantly enhance the overall user experience. Users found them confusing, distracting, or lacking impact. Some users reported losing focus on the main avatars and split their attention between the primary and notification fields. The holographic experience remained superior in terms of immersion, natural feel, and the ability to monitor all avatars simultaneously. While the notifications did have positive aspects such as assisting with avatar focus, providing directional guidance, and adding awareness, they fell short of fully bridging the gap between AR and holographic experiences.

VIII. DISCUSSION

Based on our findings, we discuss several key topics that require further exploration and in-depth analysis to better understand their implications

A. Improved Musical Collaboration with holograms

Our results clearly demonstrate the significant impact of the limited Field of View (FOV) on the quality of musical collaboration within virtual avatar-based XR experiences. The data we have gathered strongly suggests that holographic representations of avatars offer substantial advantages, including heightened enjoyment, enhanced co-presence, and swifter reaction times during collaborative musical endeavors. Drawing upon our extensive field observations and research findings, it becomes evident that effective communication plays a pivotal role in enhancing musical performance in XR environments. Musicians who experienced more seamless and immersive communication with holographic avatars consistently exhibited superior musical performance. This finding underscores the importance of addressing communication challenges within XR setups to optimize collaborative experiences. Moreover, the feedback from our participants is in line with the quantitative data, as they reported a noticeable ease and comfort while engaging with holograms compared to the alternative of interacting with virtual musicians and their own physical instruments. This sentiment further bolsters our argument that

holograms represent a superior choice for facilitating musical collaboration with avatars within the XR landscape, ultimately outclassing the capabilities of AR glasses in this particular context.

B. AR Glasses Notifications in Musical Collaboration

In previous research, Weng et al. [14] experimented with abstract notifications, specifically the blue beam notification. However, the results showed that the abstract notification did not significantly affect the drum circle setup in terms of co-presence, enjoyment, or reaction time. This outcome is attributed to the abstract notification's failure to provide relevant information about avatars. Subsequently, we introduced mini musicians, a notification depicting miniature avatars, which improved reaction time compared to no notification. Despite this improvement, neither notification enhanced the sense of co-presence and enjoyment. Participants reported that the notifications caused distraction and confusion. Consequently, we suggest that further improvements are necessary when using notifications for musical collaboration in AR glasses

C. Exploring the Potential of Social Cue Notifications for AR Glasses

In this study, we focused on investigating the effectiveness of out-of-field-of-view avatar notifications for AR glasses in jamming sessions. Our research specifically targeted social co-presence and enjoyment, rather than functional designs like navigation, task alarms, or instructions in AR glasses. The limitations imposed by FOV continue to be a challenge for state-of-the-art head-mounted XR devices. How can we inform users about out-of-FOV avatars who wish to communicate with them? We propose that there are numerous research questions that can be explored in this specific field to address this issue.

IX. LIMITATIONS AND FUTURE WORK

This study had several limitations that we aim to address in future research to further enhance our understanding and expand the possibilities of musical collaboration in XR.

A. Musical Collaborations with XR devices

Future studies could focus on exploring social notifications specifically designed for AR glasses. By integrating notifications tailored to the AR glasses platform, researchers can enhance the overall user experience and investigate the impact of social cues on musical collaboration. While our study primarily concentrated on see-through AR glasses, there are other Mixed Reality (MR) devices available that employ video pass-through technology, offering a wider field of view. Future research should consider examining the effects of these devices and compare them to see-through AR glasses to gain a comprehensive understanding of their impact on musical collaboration.

B. Intelligent Avatar Interaction

Previous research by XR Jam [4] introduced a musical collaboration system that integrates Artificial Intelligence (AI) avatars. This innovative approach allows AI avatars to participate in musical sessions, responding to the users' musical inputs. Future research could delve deeper into the potential of AI avatars to interact with the system. This involves investigating how intelligent avatars can dynamically respond to users' music playing, offering real-time feedback and interaction. The integration of AI could lead to more fluid and natural interactions, fostering a richer and more immersive environment for musical collaboration.

C. Social Scenarios for XR Explored

Our study primarily focused on the musical jamming experience, but there are other scenarios in which AR glasses could be utilized. Future work should expand its scope to explore the development of notifications for general situations that AR glasses users may encounter. This would provide a more comprehensive understanding of the potential applications and benefits of AR glasses beyond musical collaboration.

D. Evaluation Methods

While self-reported questionnaires provided valuable insights into participants' subjective experiences, future research should consider incorporating more objective data. For example, measuring participants' skin condition or utilizing brainwave monitoring devices can offer additional quantitative and physiological insights into their level of engagement, immersion, and cognitive load during the collaborative musical experience.

E. Future Study

A follow up study we would like to conduct is one that involves eye-tracking during live jam sessions. High-density eye-tracking has the potential to reveal a lot of information about the dynamics of eye-gaze during communicative tasks in AR and in live jam sessions. A comparative analysis of eye-gaze between the conditions and a normal jam session may give us in-depth insight into what is missing from the AR experience and how we can better overcome these challenges.

Additionally, interview results suggested that players experienced communication breakdowns when field of view was occluded, suggesting that visual communication is being used to recognize and correct musical mishaps. The mechanism of collective correction of the music will be studied in future experiments.

X. CONCLUSION

The objective of our study was to investigate the impact of field of view (FOV) on players' perception of communicative gestures from their partner's avatar in a musical task. Additionally, we designed various notifications for AR glasses. Our system setup involved the creation of HoloJam, an FOV-unlimited holographic installation, along with AR glasses applications featuring Mini Musicians notifications. Through our conducted study, we aimed to address our research questions. The results of our study indicated that HoloJam outperformed other setups across all measures. Additionally, the Mini Musicians notification for AR glasses contributed to reduced reaction times. Therefore, we conclude that FOV does indeed influence musical collaboration with avatars in extended reality (XR) environments. Furthermore, the use of notifications assists users in perceiving avatars more quickly. Nonetheless, there remains ample room for exploration regarding social interaction with out-of-FOV avatars in XR.

ACKNOWLEDGMENT

The authors would like to thank Gregoire Phillips, Alvin Jude, Gunilla Berndtsson, Per-Erik Brodin, Amir Gomroki, and Per Karlsson for their support and advice throughout this project. This project was supported by a grant from Ericsson Research, and by the National Science Foundation under Grant No. IIS-2040489.

REFERENCES

- [1] T. Hopkins, S. C. C. Weng, R. Vanukuru, E. A. Wenzel, A. Banic, M. D. Gross, and E. Y.-L. Do, "Ar drum circle: Real-time collaborative drumming in ar," vol. 3, 2022. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/frvir.2022.847284>
- [2] <https://jamulus.io/>.
- [3] <https://www.jacktrip.com/>.
- [4] T. Hopkins, S. C. Chuan Weng, R. Vanukuru, S. Novack, C. Tobin, E. Wenzel, A. Banic, M. Gross, and E. Y.-L. Do, "Xr jam: Design considerations for music networking with ai in xr," in *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, 2024, pp. 97–101.
- [5] A. Boem, M. Tomasetti, and L. Turchet, "Harmonizing the musical metaverse: unveiling needs, tools, and challenges from experts' point of view," 05 2024.
- [6] C. Treppowski, D. Eibich, J. Maiero, A. Marquardt, E. Kruijff, and S. Feiner, "The effect of narrow field of view and information density on visual search performance in augmented reality," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, pp. 575–584. [Online]. Available: <https://ieeexplore.ieee.org/document/8798312/>
- [7] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst, "Mini-me: An adaptive avatar for mixed reality remote collaboration," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, pp. 1–13. [Online]. Available: <https://dl.acm.org/doi/10.1145/3173574.3173620>

- [8] U. Gruenefeld, D. Lange, L. Hammer, S. Boll, and W. Heuten, "FlyingARrow: Pointing towards out-of-view objects on augmented reality devices," in *Proceedings of the 7th ACM International Symposium on Pervasive Displays*. ACM, pp. 1–6. [Online]. Available: <https://dl.acm.org/doi/10.1145/3205873.3205881>
- [9] U. Gruenefeld, T. C. Stratmann, A. E. Ali, S. Boll, and W. Heuten, "Radiallight: exploring radial peripheral leds for directional cues in head-mounted displays," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: <https://doi.org/10.1145/3229434.3229437>
- [10] L. Qian, A. Plopski, N. Navab, and P. Kazanides, "Restoring the awareness in the occluded visual field for optical see-through head-mounted displays," vol. 24, no. 11, pp. 2936–2946. [Online]. Available: <https://ieeexplore.ieee.org/document/8456571/>
- [11] K. Fan, J. Huber, S. Nanayakkara, and M. Inami, "SpiderVision: extending the human field of view for augmented awareness," in *Proceedings of the 5th Augmented Human International Conference*. ACM, pp. 1–8. [Online]. Available: <https://dl.acm.org/doi/10.1145/2582051.2582100>
- [12] Y. S. Pai, R. Hajika, K. Gupta, P. Sasikumar, and M. Billinghurst, "NeuralDrum: Perceiving brain synchronicity in XR drumming," in *SIGGRAPH Asia 2020 Technical Communications*. ACM, pp. 1–4. [Online]. Available: <https://dl.acm.org/doi/10.1145/3410700.3425434>
- [13] S. C.-C. Weng, T. Hopkins, S.-Y. Ma, C. Tobin, A. Banic, and E. Yi-Luen Do, "Investigating the effects of limited field of view on jamming experience in extended reality," in *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2023, pp. 493–495.
- [14] S. C. Weng, T. Hopkins, R. Vanukuru, C. Tobin, A. Banic, D. Leithinger, and E. Y. Do, "How field of view affects awareness of an avatar during a musical task in augmented reality," in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2023, pp. 633–634.
- [15] L. Turchet, R. Hamilton, and A. Çamci, "Music in extended realities," *IEEE Access*, vol. 9, pp. 15 810–15 832, 2021.
- [16] R. Schlagowski, D. Nazarenko, Y. Can, K. Gupta, S. Mertes, M. Billinghurst, and E. André, "Wish you were here: Mental and physiological effects of remote music collaboration in mixed reality," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3544548.3581162>
- [17] R. Berry, M. Makino, N. Hikawa, and M. Suzuki, "The augmented composer project: the music table," in *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2003. *Proceedings.*, 2003, pp. 338–339.
- [18] S. Jordà, G. Geiger, M. Alonso, and M. Kaltenbrunner, "The reactable: Exploring the synergy between live music performance and tabletop tangible interfaces," in *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, ser. TEI '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 139–146. [Online]. Available: <https://doi.org/10.1145/1226969.1226998>
- [19] N. Bryan-Kinns, "Daisyphone: The design and impact of a novel environment for remote group music improvisation," in *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, ser. DIS '04. New York, NY, USA: Association for Computing Machinery, 2004, p. 135–144. [Online]. Available: <https://doi.org/10.1145/1013115.1013135>
- [20] L. Men and N. Bryan-Kinns, "Lemo: Exploring virtual space for collaborative creativity," in *Proceedings of the 2019 on Creativity and Cognition*, ser. C&C '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 71–82. [Online]. Available: <https://doi.org/10.1145/3325480.3325495>
- [21] p. cairns, a. hunt, d. johnston, j. cooper, b. lee, h. daffern, and g. kearney, "evaluation of metaverse music performance with bbc maida vale recording studios," *journal of the audio engineering society*, vol. 71, no. 6, pp. 313–325, june 2023.
- [22] E. Costanza, S. A. Inverso, E. Pavlov, R. Allen, and P. Maes, "Eye-q: Eyeglass peripheral display for subtle intimate notifications," in *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 211–218. [Online]. Available: <https://doi.org/10.1145/1152215.1152261>
- [23] S. Matsuzoe, S. Jiang, M. Ueki, and K. Okabayashi, "Intuitive visualization method for locating off-screen objects inspired by motion perception in peripheral vision," in *Proceedings of the 8th Augmented Human International Conference*, ser. AH '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: <https://doi.org/10.1145/3041164.3041198>
- [24] L. Plabst, S. Oberdörfer, F. R. Ortega, and F. Niebling, "Push the red button: Comparing notification placement with augmented and non-augmented tasks in ar," in *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*, ser. SUI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: <https://doi.org/10.1145/3565970.3567701>
- [25] I. Chaturvedi, F. H. Bijarbooneh, T. Braud, and P. Hui, "Peripheral vision: A new killer app for smart glasses," in *Proceedings of the 24th International Conference on Intelligent User Interfaces*, ser. IUI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 625–636. [Online]. Available: <https://doi.org/10.1145/3301275.3302263>
- [26] R. Rzayev, S. Korbely, M. Maul, A. Schark, V. Schwind, and N. Henze, "Effects of position and alignment of notifications on ar glasses during social interaction," in *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, ser. NordiCHI '20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: <https://doi.org/10.1145/3419249.3420095>
- [27] F. Biocca and C. Harms, "Defining and measuring social presence: Contribution to the networked minds theory and measure," *Proceedings of PRESENCE*, vol. 2002, pp. 7–36, 2002.
- [28] D. Görlich, "Societal xr—a vision paper," *ParadigmPlus*, vol. 3, no. 2, pp. 1–10, 2022.
- [29] J. Hwang, G. JounghyunKim, and H. Kim, "Incorporating co-presence in distributed virtual music environment," p. 6.
- [30] N. Osmers, M. Prilla, O. Blunk, G. George Brown, M. Janßen, and N. Kahl, "The role of social presence for cooperation in augmented reality on head mounted devices: A literature review," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: <https://doi.org/10.1145/3411764.3445633>
- [31] J. D. Hart, T. Piumsomboon, G. A. Lee, R. T. Smith, and M. Billinghurst, "Manipulating avatars for enhanced communication in extended reality," in *2021 IEEE International Conference on Intelligent Reality (ICIR)*, May 2021, pp. 9–16.
- [32] M. Kurzweg and K. Wolf, "Body language of avatars in vr meetings as communication status cue: Recommendations for interaction design and implementation," *i-com*, vol. 21, no. 1, pp. 175–201, 2022. [Online]. Available: <https://doi.org/10.1515/icom-2021-0038>
- [33] H. Bai, P. Sasikumar, J. Yang, and M. Billinghurst, "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*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–13. [Online]. Available: <https://doi.org/10.1145/3313831.3376550>
- [34] S. Kimmel, F. Jung, A. Matvienko, W. Heuten, and S. Boll, "Let's face it: Influence of facial expressions on social presence in collaborative virtual reality," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: <https://doi.org/10.1145/3544548.3580707>
- [35] S. Greenwald, Z. Wang, M. Funk, and P. Maes, "Investigating social presence and communication with embodied avatars in room-scale virtual reality," 06 2017, pp. 75–90.
- [36] J. Blatterste, B. Strenge, P. Renner, T. Pfeiffer, and K. Essig, "Comparing conventional and augmented reality instructions for manual assembly tasks," in *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments*, ser. PETRA '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 75–82. [Online]. Available: <https://doi.org/10.1145/3056540.3056547>
- [37] Y. Shin and G. J. Kim, "Leveraging vr techniques for efficient exploration and interaction in large and complex ar space with clipped and small fov ar display," in *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: <https://doi.org/10.1145/3562939.3565613>

- [38] "Datavyu," <https://datavyu.org>, accessed: June 12, 2023.
- [39] M. Trajkova, M. Deshpande, A. Knowlton, C. Monden, D. Long, and B. Magerko, "Ai meets holographic pepper's ghost: A co-creative public dance experience," in *Companion Publication of the 2023 ACM Designing Interactive Systems Conference*, ser. DIS '23 Companion. New York, NY, USA: Association for Computing Machinery, 2023, p. 274–278. [Online]. Available: <https://doi.org/10.1145/3563703.3596658>
- [40] S. Lee, D. Kim, and J. Lee, "The effects of viewing formats and song genres on audience experiences in virtual avatar concerts," in *Proceedings of the 31st ACM International Conference on Multimedia*, ser. MM '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 6978–6988. [Online]. Available: <https://doi.org/10.1145/3581783.3612094>
- [41] M. Hamanaka, "Melody slot machine: A controllable holographic virtual performer," in *Proceedings of the 27th ACM International Conference on Multimedia*, ser. MM '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 2468–2477. [Online]. Available: <https://doi.org/10.1145/3343031.3350863>
- [42] M. Slater, C. Cabrera, G. Senel, D. Banakou, A. Beacco, R. Oliva, and J. Gallego, "The sentiment of a virtual rock concert," *Virtual Reality*, vol. 27, no. 2, pp. 651–675, Jun 2023. [Online]. Available: <https://doi.org/10.1007/s10055-022-00685-9>