



OCTOBER, 2023

MRMAC : Mixed Reality Multi-user Asymmetric Collaboration

 **Faisal Zaman**

 Craig Anslow

 Andrew Chalmers

 Taehyun Rhee

CMIC | Computational
Media
Innovation
Centre



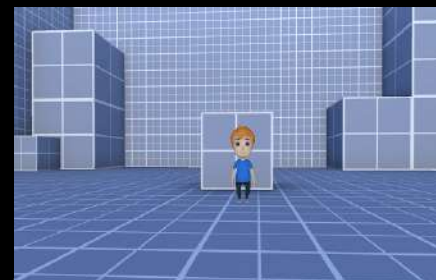
VICTORIA UNIVERSITY OF
WELLINGTON
TE HERENGA WAKA



 IEEE
COMPUTER
SOCIETY

 IEEE
vgtc

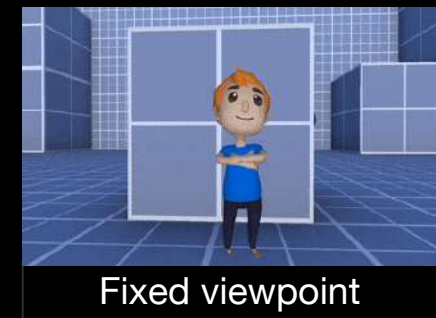
Collaborative Tools



Limited Field of view

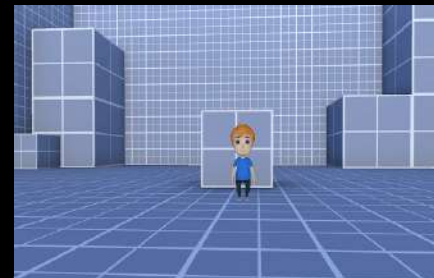


Barely see others' body language & gaze



Fixed viewpoint

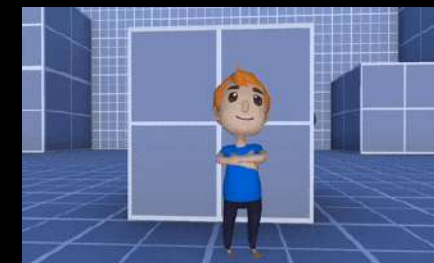
Collaborative Tools



Limited Field of view



Barely see others' body language & gaze



Fixed viewpoint

Collaborative XR Tools



Project starline, *Lawrence et al. 2021*



Microsoft Remote Assist



Spatial.io



Microsoft Mesh

Collaborative XR Tools



Spatial.io



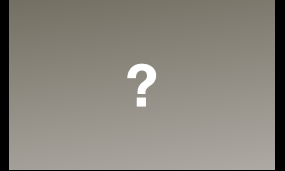
Project starline, Lawrence et al. 2021



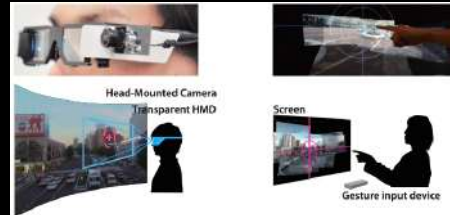
Microsoft Remote Assist



Microsoft Mesh



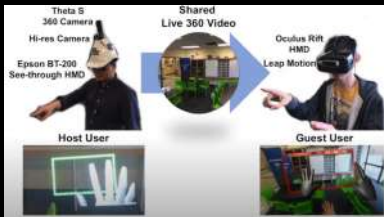
Related Work



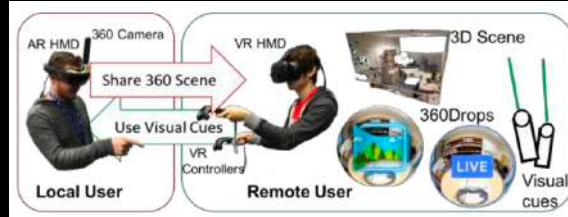
[S. Kasahara and J. Rekimoto 2014]



[S. Nagai et al. 2015]



[Lee et al. 2017]



[Teo et al. 2017]



[Lee et al. 2020]

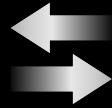


[Rhee et al. 2020]

Only One-One & One-Many

No Audio-Visual Sync

Contributions



Mixed Reality Multi-user Asymmetric Collaboration (MRMAC) system that enables multiple remote users to virtually teleport into a real-world task space to collaborate with local users.

MRMAC

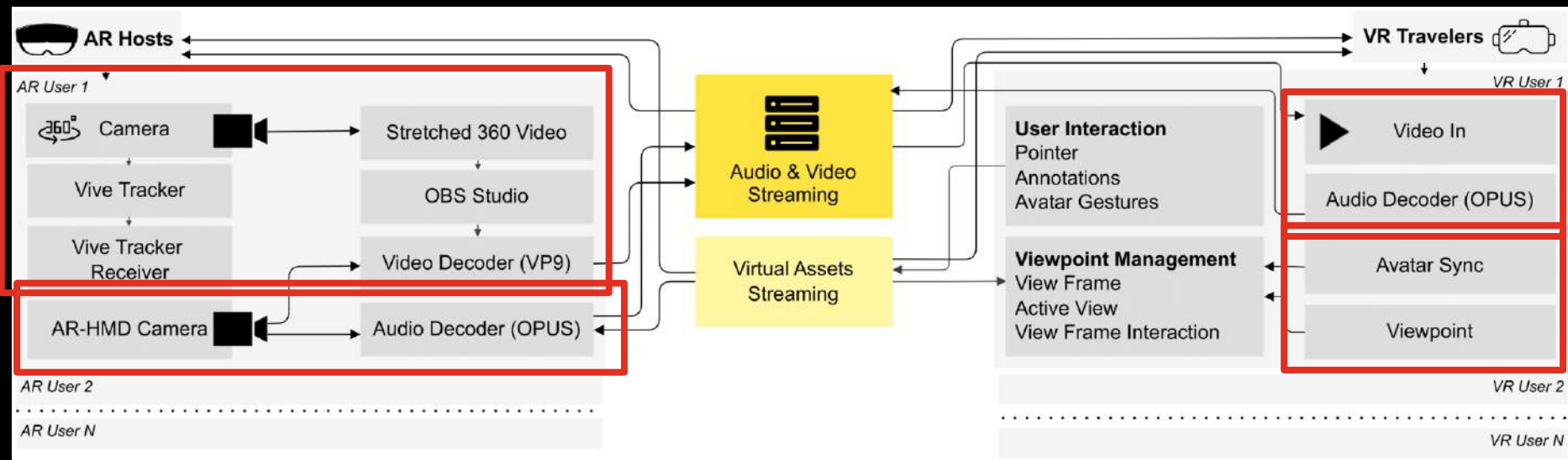
Design & Implementation

- Design concept for multi-user asymmetric remote collaboration
- Bidirectional face-to-face communication
- Synchronised audio-visual communication
- Client-server architecture for avatars, view controls, and asset streaming

Evaluation

- System evaluation
 - scalability and latency
- User evaluation
 - communication and awareness
 - compared against two baseline conditions

Architecture



Communication Cues

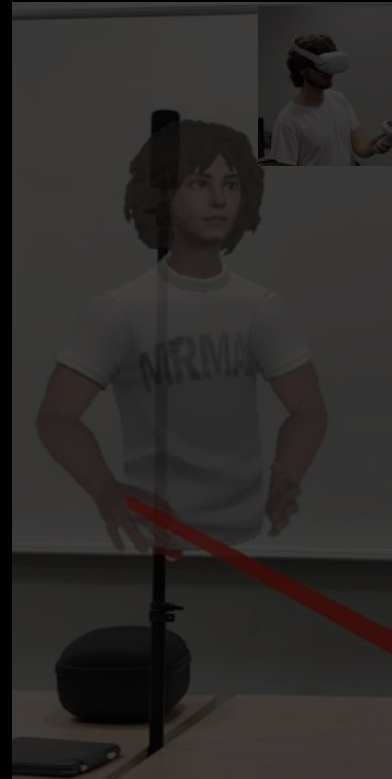
Viewpoint Sharing



Multiple Avatar Control



Personalised Avatar



Spatial Audio



Communication Cues

Viewpoint Sharing



Vicarious: Context-aware Viewpoints Selection for Mixed Reality Collaboration

Faisal Zaman
Computational Media Innovation
Centre, Victoria University of
Wellington
New Zealand
faisal.zaman@vuw.ac.nz

Craig Anslow
School of Engineering and Computer
Science, Victoria University of
Wellington
New Zealand
craig.anslow@vuw.ac.nz

Taehyun Rhee
Computational Media Innovation
Centre, Victoria University of
Wellington
New Zealand
taehyun.rhee@vuw.ac.nz



Figure 1: Two users in a mixed reality environment.

ABSTRACT

Mixed reality environments, particularly those that support collaborative work, often suffer from a lack of context-awareness. This paper presents a system called Vicarious, which provides a context-aware selection of viewpoints for users in a mixed reality environment. Vicarious is designed to be used in a mixed reality environment where multiple users are present. It allows users to select a viewpoint that is relevant to their current task, thereby improving their understanding of the environment and their ability to collaborate effectively. The system is evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task.

Keywords

Mixed reality, Viewpoint selection, Context-awareness, Collaboration, Multi-user systems

1. INTRODUCTION

Mixed reality environments, particularly those that support collaborative work, often suffer from a lack of context-awareness. This paper presents a system called Vicarious, which provides a context-aware selection of viewpoints for users in a mixed reality environment. Vicarious is designed to be used in a mixed reality environment where multiple users are present. It allows users to select a viewpoint that is relevant to their current task, thereby improving their understanding of the environment and their ability to collaborate effectively. The system is evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task.

2. RELATED WORK

Viewpoint selection in mixed reality environments is a challenging task. It requires the system to understand the user's current task and to select a viewpoint that is relevant to that task. This is often done by using a series of heuristics, such as the user's current position and the location of the objects in the environment. However, these heuristics are often not sufficient to select the best viewpoint for a given task. This paper presents a system called Vicarious, which provides a context-aware selection of viewpoints for users in a mixed reality environment. Vicarious is designed to be used in a mixed reality environment where multiple users are present. It allows users to select a viewpoint that is relevant to their current task, thereby improving their understanding of the environment and their ability to collaborate effectively. The system is evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task.

3. SYSTEM ARCHITECTURE

The system architecture is shown in Figure 2. It consists of a client device, a server, and a database. The client device is responsible for capturing the user's current task and for sending this information to the server. The server is responsible for selecting a viewpoint that is relevant to the user's current task and for sending this information back to the client device. The database is responsible for storing the user's current task and for providing this information to the server. The system is evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task.

4. EVALUATION

The system is evaluated using a series of experiments. The first experiment is designed to evaluate the system's ability to select a viewpoint that is relevant to the user's current task. The second experiment is designed to evaluate the system's ability to improve users' performance in a collaborative task. The third experiment is designed to evaluate the system's ability to improve users' understanding of the environment. The results of the experiments show that the system significantly improves users' performance in a collaborative task and that it significantly improves users' understanding of the environment. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task.

5. CONCLUSION

This paper presents a system called Vicarious, which provides a context-aware selection of viewpoints for users in a mixed reality environment. Vicarious is designed to be used in a mixed reality environment where multiple users are present. It allows users to select a viewpoint that is relevant to their current task, thereby improving their understanding of the environment and their ability to collaborate effectively. The system is evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task. The system is also evaluated using a series of experiments, and the results show that it significantly improves users' performance in a collaborative task.

6. ACKNOWLEDGMENTS

This work was supported by the New Zealand Research Council.

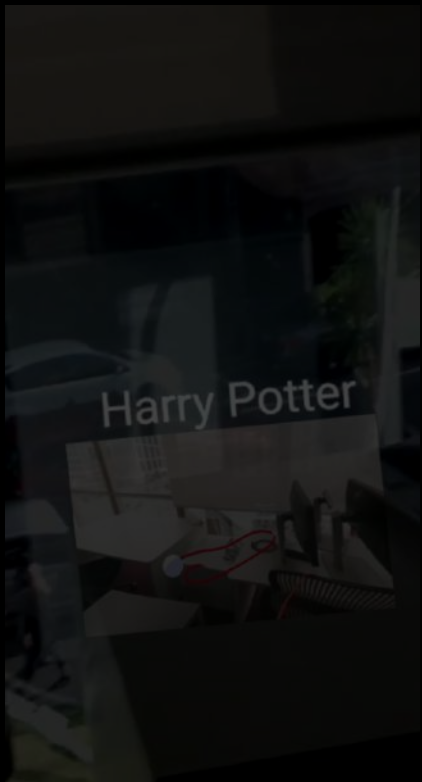
7. REFERENCES

[1] F. Zaman, C. Anslow, and T. Rhee, "Vicarious: Context-aware Viewpoints Selection for Mixed Reality Collaboration," in *ACM VRST*, 2023.

F. Zaman, C. Anslow, T. Rhee.
Vicarious: Context-aware Viewpoints Selection for Mixed Reality Collaboration.
ACM VRST, 2023.

Communication Cues

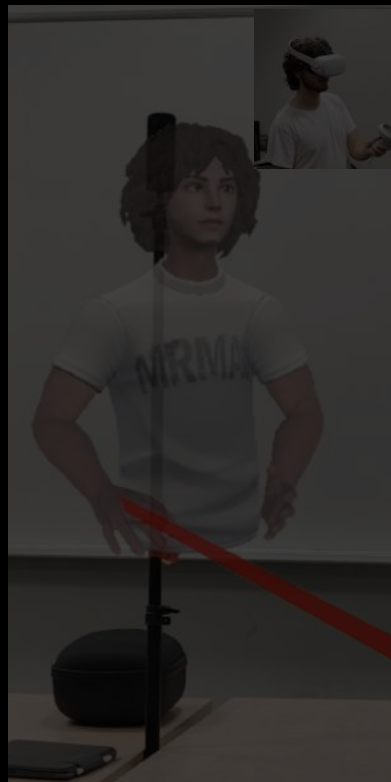
Viewpoint Sharing



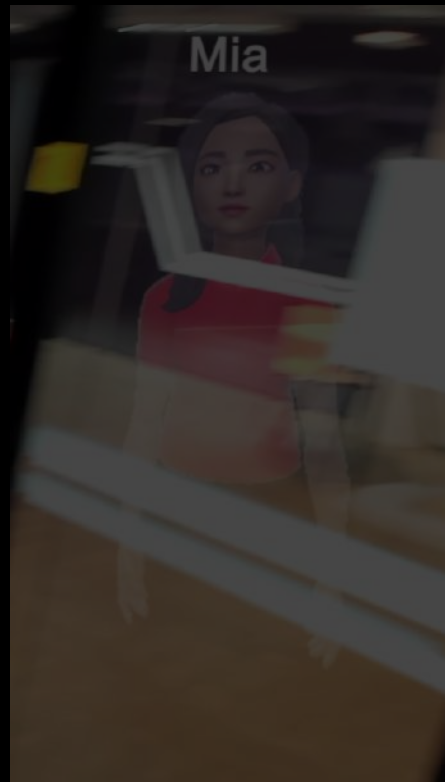
Multiple Avatar Control



Personalised Avatar



Spatial Audio



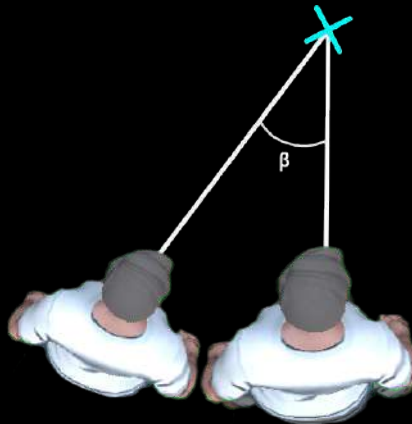
Communication Cues

Multiple Avatar Control



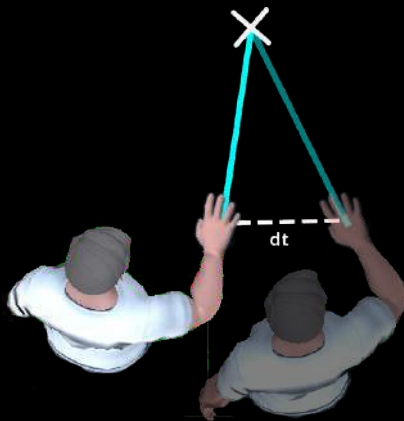
Communication Cues

Multiple Avatar Control



Communication Cues

Multiple Avatar Control



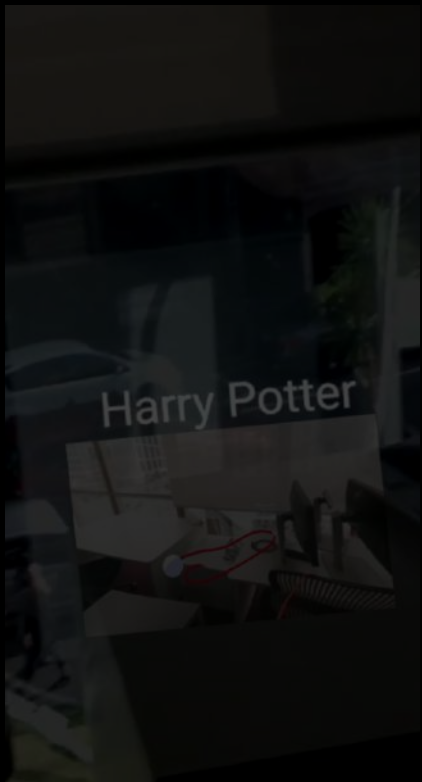
Communication Cues

Multiple Avatar Control



Communication Cues

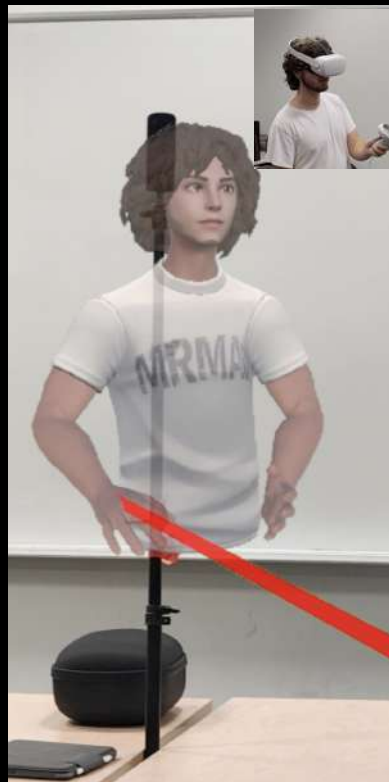
Viewpoint Sharing



Multiple Avatar Control



Personalised Avatar

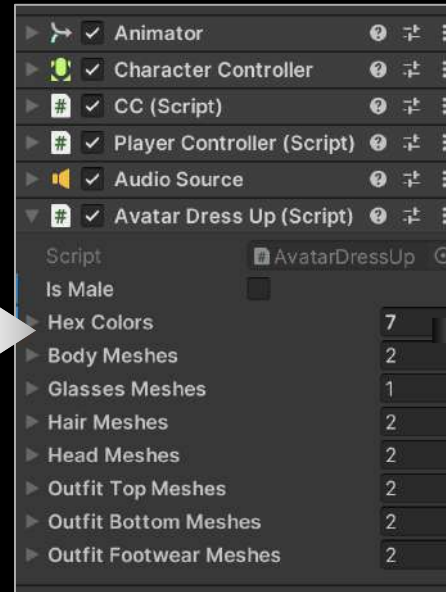


Spatial Audio



Communication Cues

Personalised Avatar



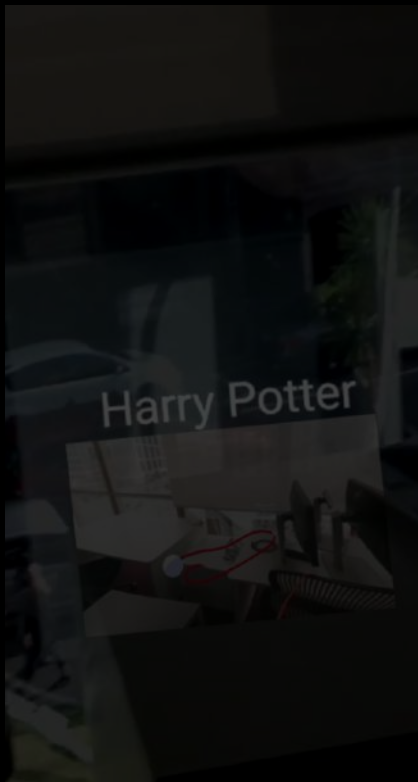
READY
PLAYER
ME



HEADSHOT2

Communication Cues

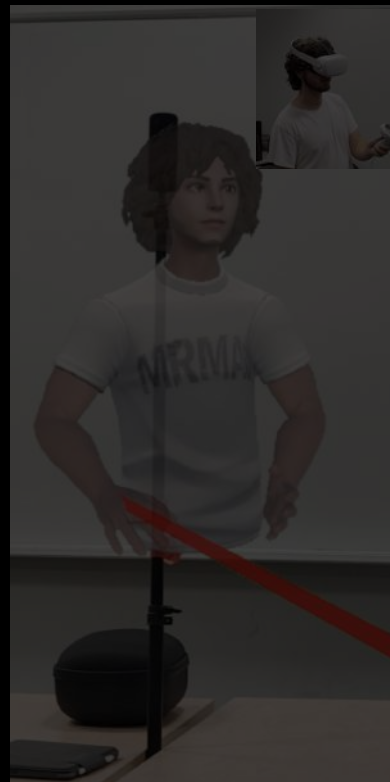
Viewpoint Sharing



Multiple Avatar Control



Personalised Avatar



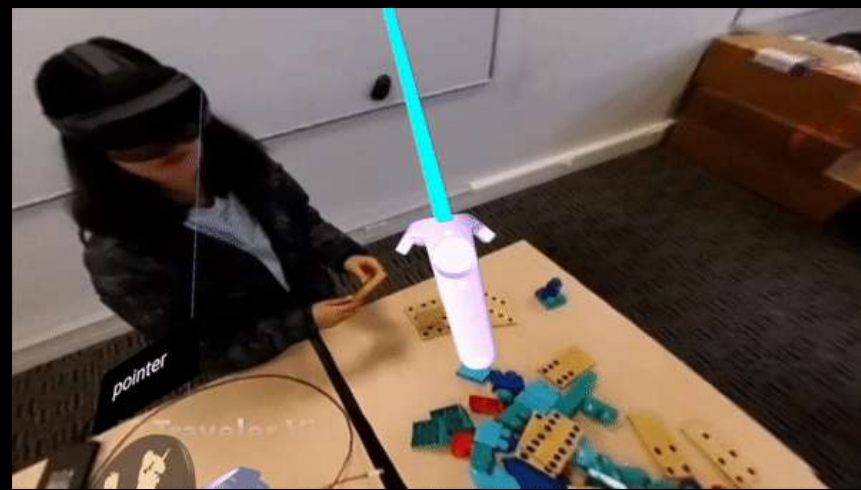
Spatial Audio



Interaction Cues



3D Pointer



Annotation

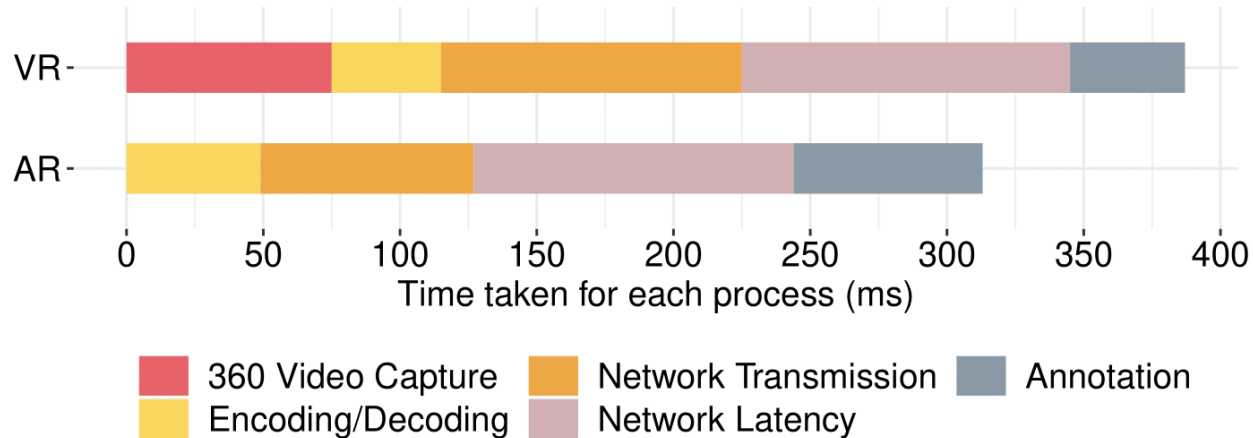
evaluation

Evaluation

System Evaluation: Latency

Average end-to-end latency

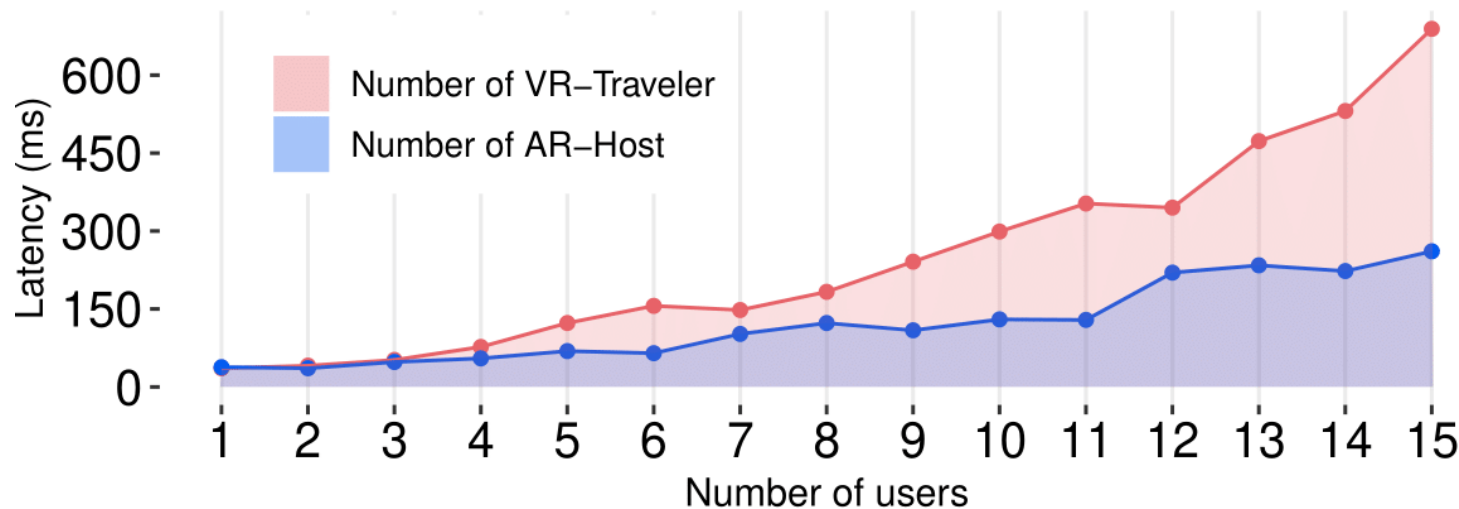
- Changes in the 360-degree video to show up on the local computer and in the VR/AR window.



System Evaluation: Scalability

Network latency as the number of users increases

- average frame rate = video streaming 30FPS
- audio streaming 44.1kHz
- rendering time of 60 ± 10 FPS.



User Evaluation

36 participants

18—81 years

$\mu = 30.83$, $\sigma = 14.10$

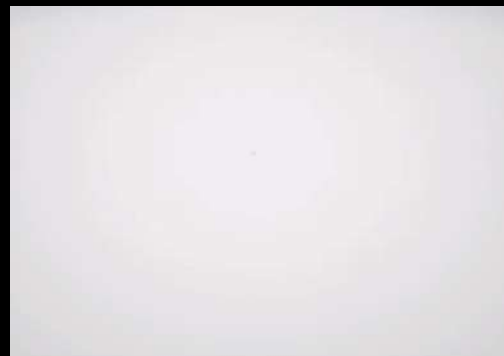
3×2 mixed factorial design

between-subjects factor:
local vs. remote

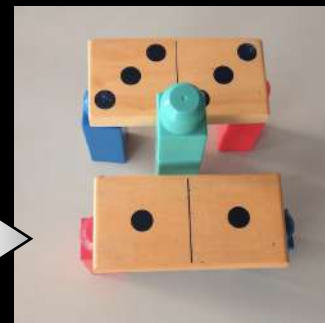
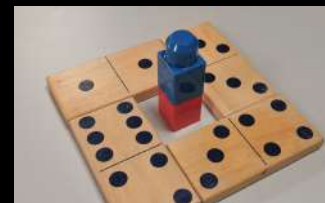
within-subjects factor:
C1 vs. C2 vs. C3

VR Users

AR Users



Virtual replica



Physical model

User Evaluation

3×2 mixed factorial design

between-subjects factor:
local vs. remote

within-subjects factor:
C1 vs. C2 vs. C3

1

Condition 1
Conventional video
with 2D annotation

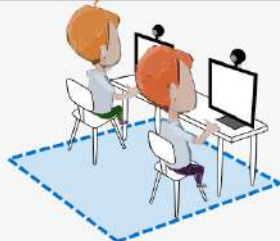
2

Condition 2
360 Video without
augmented visual cues

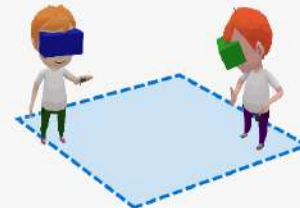
3

Condition 3
360 video with
augmented visual cues

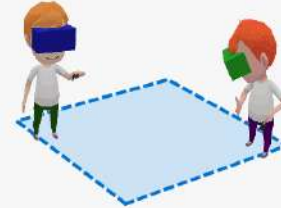
VR-Travelers



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
Mouse for 2D Annotation
Representation:
Appears in 2D video

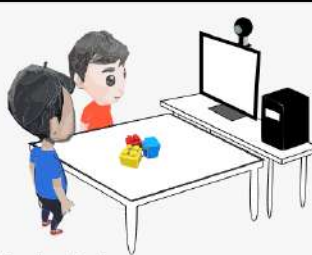


VR HMD (Viewing 360 Video)
Camera Control by VR HMD
N/A
No Visual, Voice Only

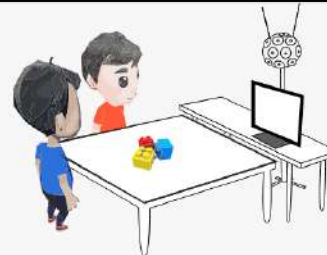


VR HMD (Viewing 360 Video)
Camera Control by VR HMD
VR Controller
3D Avatar

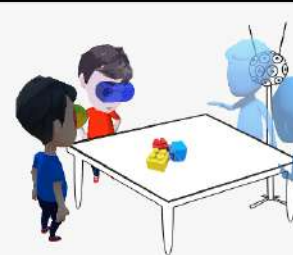
AR-Hosts



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
N/A
Representation:
Appears in 2D video



Desktop Monitor
N/A
N/A
Appears in 360-degree Video



AR HMD
N/A
N/A
Appears in 360-degree Video

User Evaluation

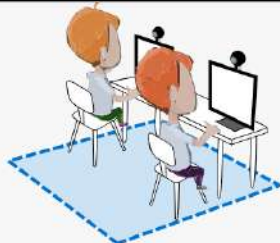
Condition 1



1

Condition 1
Conventional video
with 2D annotation

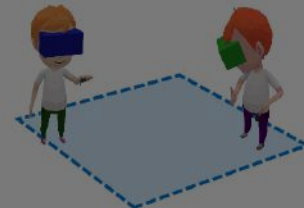
VR-Travelers



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
Mouse for 2D Annotation
Representation:
Appears in 2D video

2

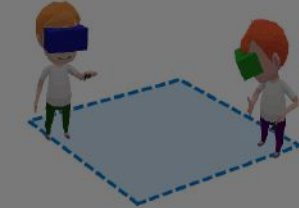
Condition 2
360 Video without
augmented visual cues



VR HMD (Viewing 360 Video)
Camera Control by VR HMD
N/A
No Visual, Voice Only

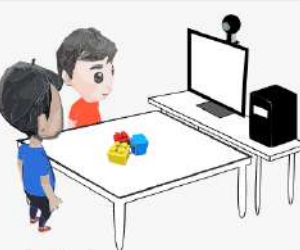
3

Condition 3
360 video with
augmented visual cues

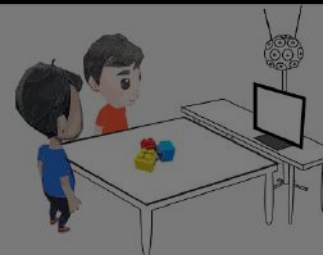


VR HMD (Viewing 360 Video)
Camera Control by VR HMD
VR Controller
3D Avatar

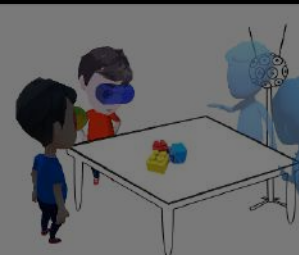
AR-Hosts



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
N/A
Representation:
Appears in 2D video



Desktop Monitor
N/A
N/A
Appears in 360-degree Video



AR HMD
N/A
N/A
Appears in 360-degree Video

User Evaluation

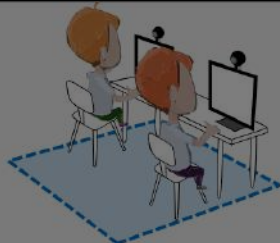
Condition 2



1

Condition 1
Conventional video
with 2D annotation

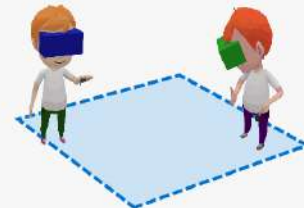
VR-Travelers



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
Mouse for 2D Annotation
Representation:
Appears in 2D video

2

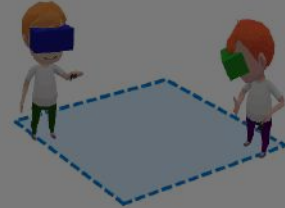
Condition 2
360 Video without
augmented visual cues



VR HMD (Viewing 360 Video)
Camera Control by VR HMD
N/A
No Visual, Voice Only

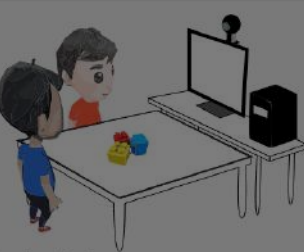
3

Condition 3
360 video with
augmented visual cues

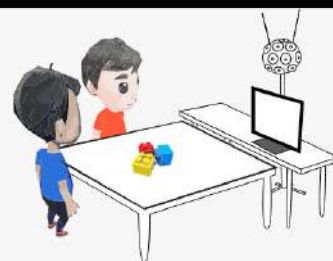


VR HMD (Viewing 360 Video)
Camera Control by VR HMD
VR Controller
3D Avatar

AR-Hosts



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
N/A
Representation:
Appears in 2D video



Desktop Monitor
N/A
N/A
Appears in 360-degree Video



AR HMD
N/A
N/A
Appears in 360-degree Video

User Evaluation

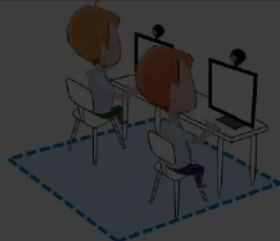
Condition 3



1

Condition 1
Conventional video
with 2D annotation

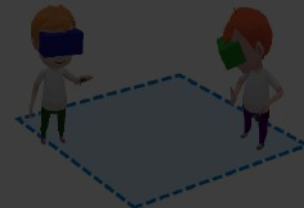
VR-Travelers



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
Mouse for 2D Annotation
Representation:
Appears in 2D video

2

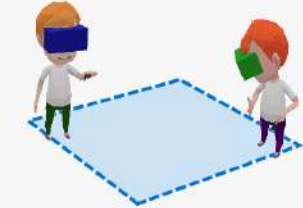
Condition 2
360 Video without
augmented visual cues



VR HMD (Viewing 360 Video)
Camera Control by VR HMD
N/A
No Visual, Voice Only

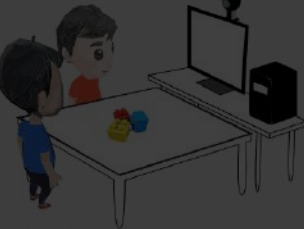
3

Condition 3
360 video with
augmented visual cues

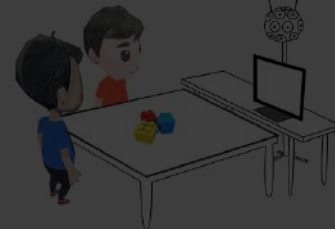


VR HMD (Viewing 360 Video)
Camera Control by VR HMD
VR Controller
3D Avatar

AR-Hosts



Viewing Device:
Desktop Monitor
Camera control:
N/A
Interaction:
N/A
Representation:
Appears in 2D video



Desktop Monitor
N/A
N/A
Appears in 360-degree Video



AR HMD
N/A
N/A
Appears in 360-degree Video

Results

Results

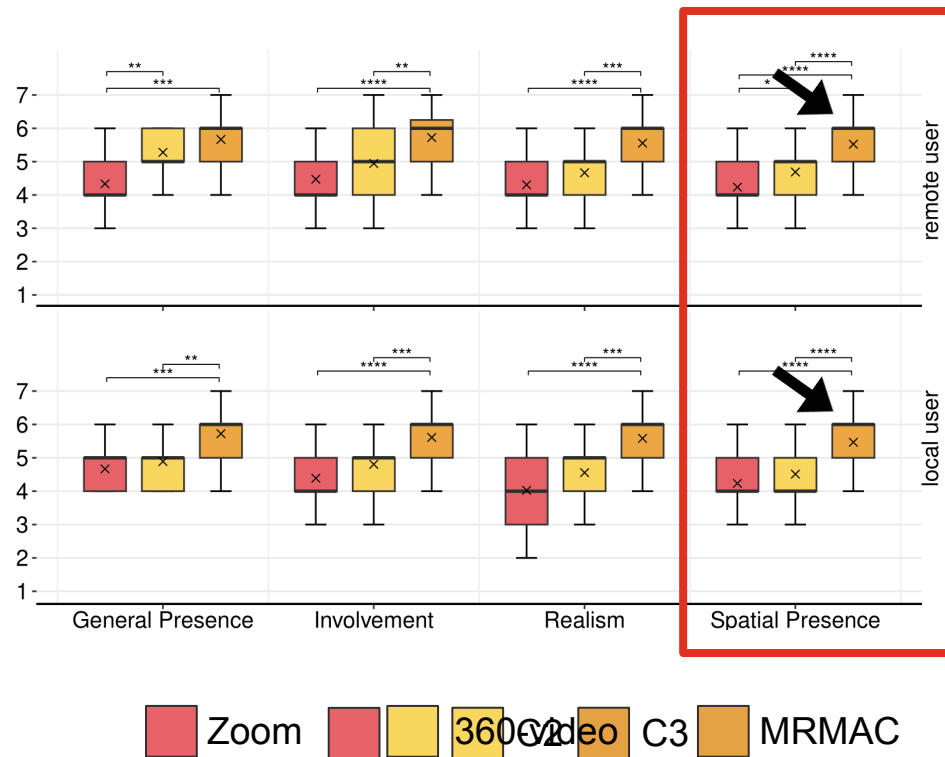
Results: Spatial Presence

H1: Spatial presence would be significantly higher in MRMAC.

Accepted

Justification:

Our system allow users to see each other and explore the real environment of users, social presence and spatial presence should be higher between AR and VR users.



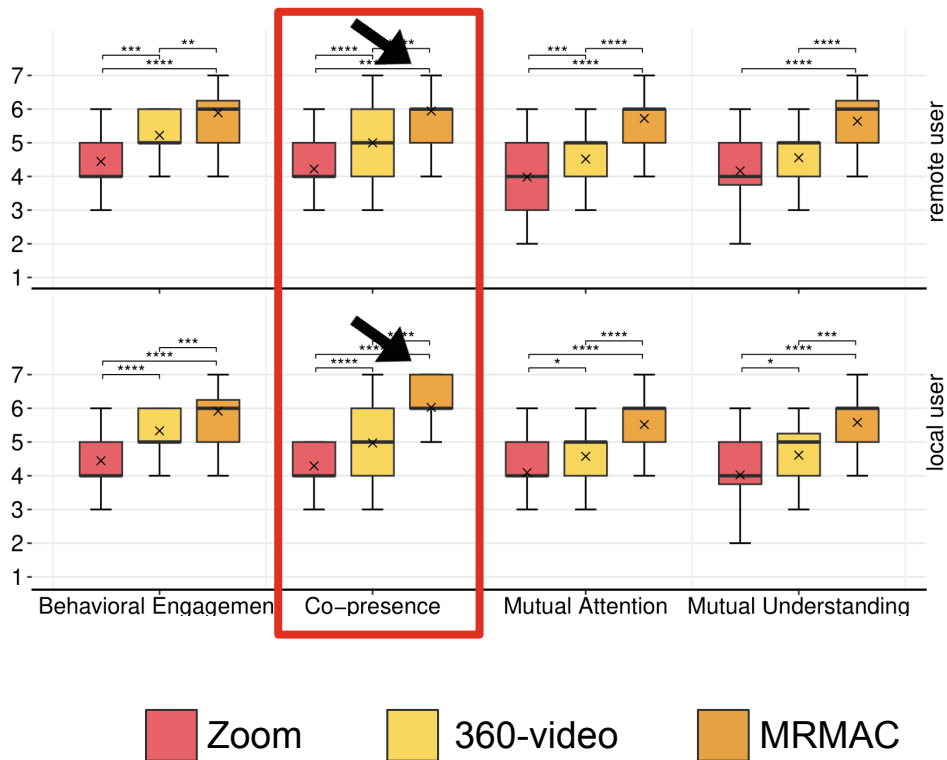
Results: Social Presence

H1: Social presence would be significantly higher in MRMAC.

Accepted

Justification:

Our system allow users to see each other and talk to each other.



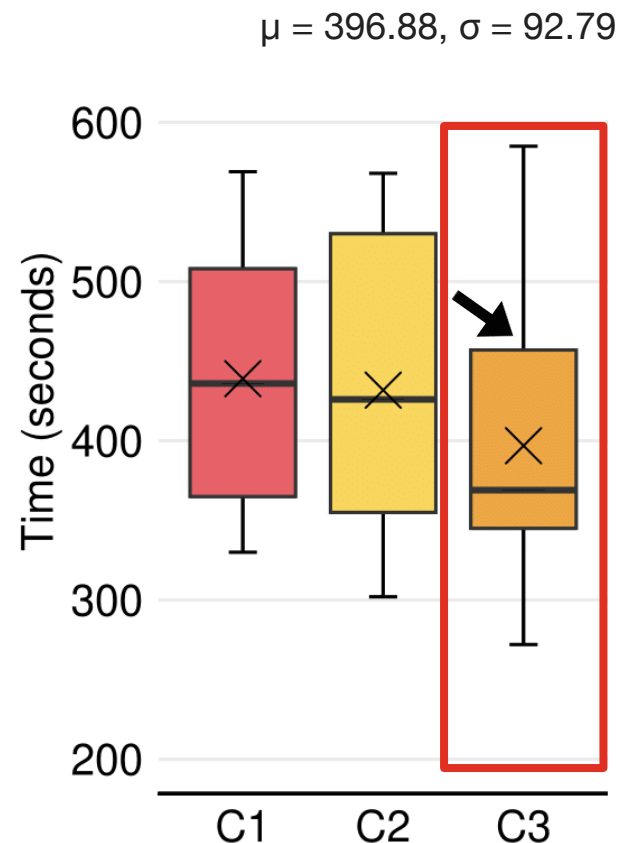
Results: Task Completion Time

H2: With MRMAC participants will complete the task faster.

Partially Accepted

Justification:

Available communication and awareness tools in MRMAC made it significantly easier for remote and local users to complete the task faster.



Results: Work Load (TLX)

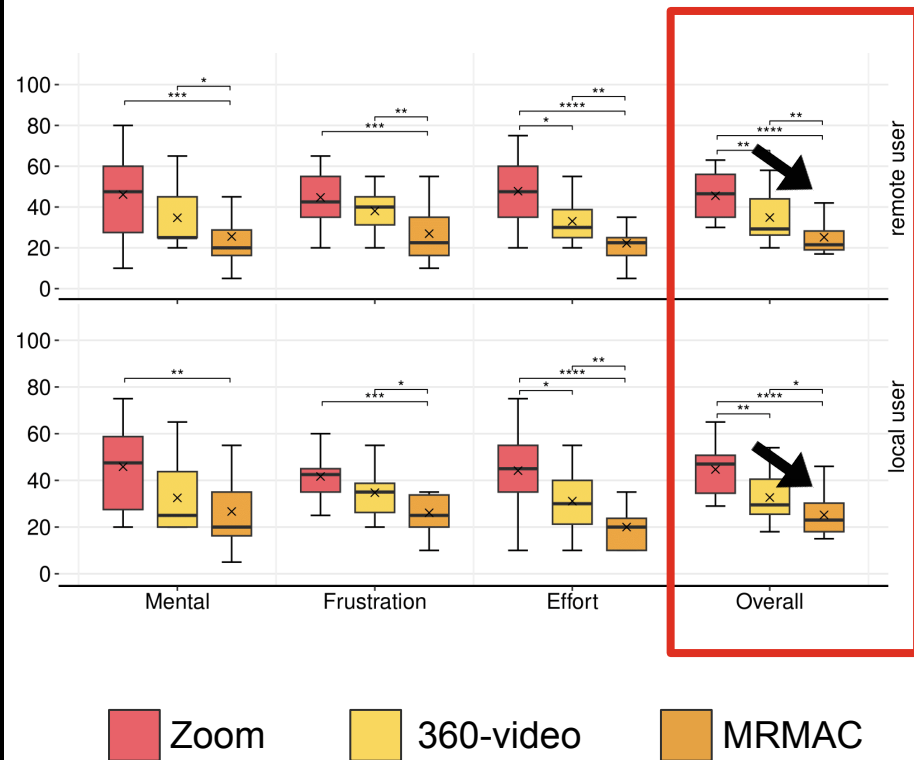
H2: With MRMAC participants will have less workload.

Accepted

Measured by: TLX

Justification:

Available communication and awareness tools in MRMAC made it significantly easier for remote and local users to complete the task faster.

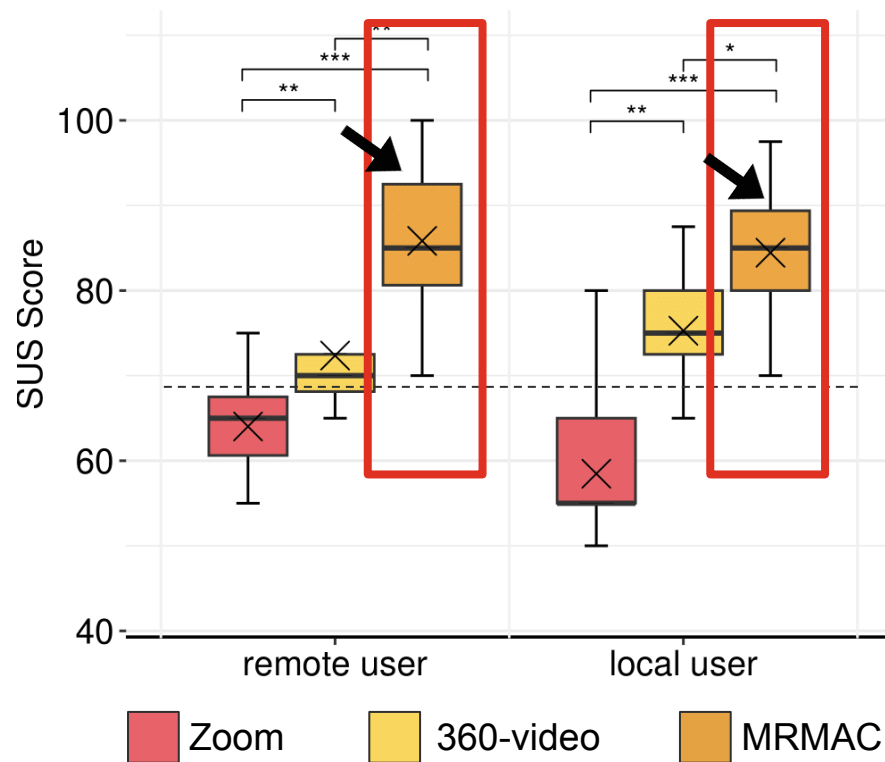


Results: Usability

H3: System usability would be significantly higher in MRMAC.

Accepted

Measured by: SUS



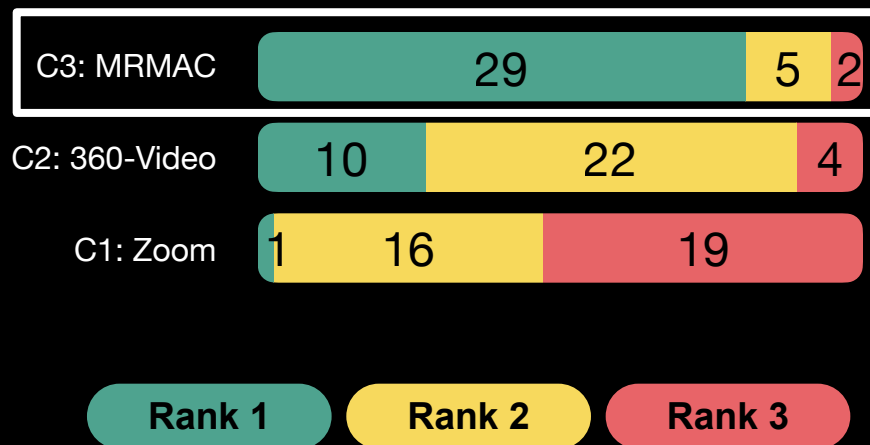
Results: Preference

H4: Participants would prefer MRMAC more.

[...] live-streaming the physical environment and blending 3D virtual assets really amps up the collaboration experience

[...] it feels like we're all in the same room. I can hear everyone's voices coming from different directions, it even more realistic

Q. Which condition did you prefer overall?



Limitation and Future Work

Explore more capture modalities

- Integrate depth information.
- Teleoperate robots.

System adaptability in challenging environment

Conduct more in-depth user studies.

conclusion

Conclusion

Summary

Fully bidirectional asymmetric collaboration system

- Real-time collaboration
- Audio-visual synchronisation
- Virtual assets blending

System evaluation

- Reasonably Scalable
- Low-latency

User evaluation

- MRMAC *performed* collaborative tasks faster
- MRMAC *preferred* over conventional 2D and traditional 360

faisal.zaman@vuw.ac.nz
[jquery404.github.io](https://github.com/jquery404)



Faisal Zaman



Craig Anslow



Andrew Chalmers



Taehyun Rhee



Acknowledgements

