MRTransformer: Transforming Avatar Non-verbal Behavior for Remote MR Collaboration in Incongruent Spaces

Cheng Yao Wang* Cornell University

Hyunju Kim† Cornell University

Payod Panda‡ Microsoft Research

Eyal Ofek§ Microsoft Research

Mar Gonzalez Franco¶ Microsoft Research

Andrea Stevenson Won‖ Cornell University

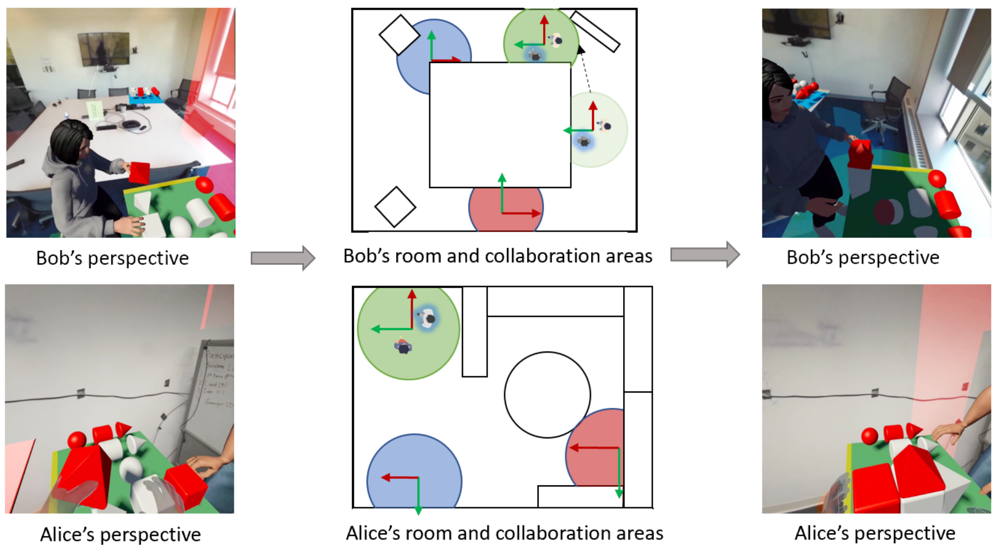

Figure 1: Demonstrating MRTransformer's dynamic management of collaboration spaces in real-time. (Left) Alice and Bob collaborate within a designated mapped area, (Middle) Bob independently adjusts the collaboration area during the activity, and (Right) The movements of both Alice's and Bob's avatars are seamlessly preserved as the collaboration area is adjusted.

Abstract

Avatar-mediated remote MR collaboration allows users in different spaces to interact as if they were together. However, directly applying a user's motion to an avatar in incongruent spaces leads to ambiguous and error-prone communication. This paper introduces *MRTransformer*, a technique enabling dynamic MR collaboration across dissimilar spaces. By adapting transformations to user movements, MRTransformer preserves non-verbal cues and spatial context. It also allows flexible management of collaboration areas and remote object visualization, enhancing remote collaborations. A user study evaluated MRTransformer's effectiveness in preserving non-verbal cues and spatial awareness, and examined social presence and privacy concerns. Findings offer implications for future remote MR collaboration research and design.

Index Terms: MR Collaboration, Avatar Movement Transformations, Non-verbal Behavior

*e-mail:chengyao.wang@jpmchase.com
†e-mail: hk724@cornell.edu
‡payod.panda@microsoft.com
§e-mail: eyal.ofek@gmail.com
¶margon@google.com
‖asw248@cornell.edu

1 Introduction

Remote work has sharply risen in recent years, spurred by technological innovations enabling remote interactions. Recent advancements in MR technology have enabled 3D avatar-based remote collaboration, allowing local users to interact with remote users through virtual avatars representing them as if they were in the same space [10, 14, 12]. However, one-to-one mapping of a remote user's motion to their avatar is inadequate when the two spaces differ in shape and object arrangements.

Previous approaches to address this issue include warping environments or finding optimal alignments, but these often result in limited and non-reconfigurable spaces [7, 4]. Recent research has proposed mapping relevant task surfaces or areas rather than aligning entire spaces [1, 2]. However, these methods still face challenges such as occlusion of remote avatars by surrounding objects and static mapped areas [8, 15].

Our research addresses two critical challenges: adapting task surfaces dynamically during collaboration and enhancing spatial awareness around these surfaces. We introduce MRTransformer, a novel system designed for dynamic collaborative environments. MRTransformer features Dynamic Collaboration Areas, which can be moved during collaboration, allowing avatars' movements to be adjusted in real-time. This ensures coherent interactions across various physical layouts and enhances spatial awareness by visualizing occlusions, continuously updating these visualizations as collaboration areas are adjusted [9, 3, 11].

To investigate MRTransformer's feasibility and its impact on social presence, perceptions, and privacy concerns, we conducted a

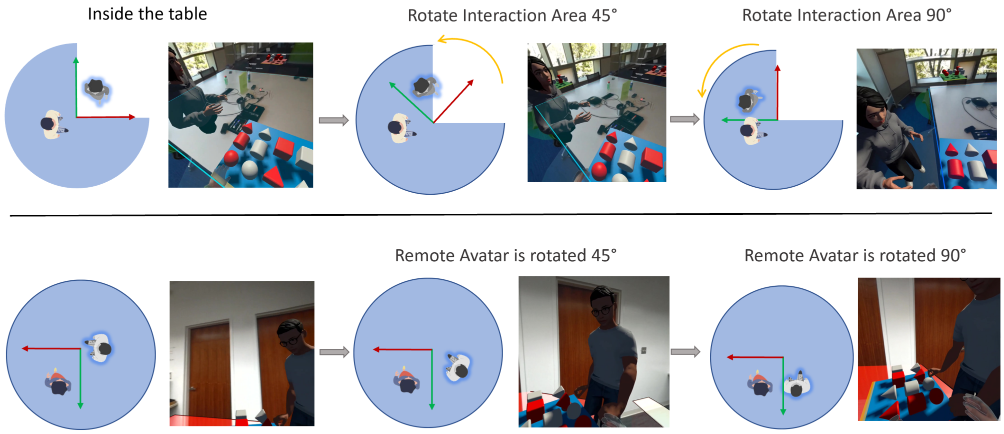

Figure 2: Illustration of rotating a Collaboration Area to move a remote avatar out of occlusion.

user study comparing it with a baseline MR collaboration system using one-to-one motion mapping. Participants preferred MRTransformer for its enhanced efficiency and reduced cognitive workload. They appreciated the flexibility to adjust and move task areas dynamically, manage avatar occlusions, and reposition after spontaneous events [6, 5, 13].

Our key contributions include the introduction of MRTransformer, a user study assessing its feasibility and impact, and a discussion of the implications for future research on remote MR collaboration in incongruent spaces.

2 MRTransformer System Design

2.1 Dynamic Collaboration Areas

MRTransformer introduces ***Dynamic Collaboration Areas***, distinct interaction zones linked to virtual or physical objects. Unlike fixed task surfaces, these areas can be moved during collaboration. Each area acts as an independent local coordinate system, allowing avatars to be visualized in remote spaces, preserving spatial and nonverbal cues. When a user is inside a Collaboration Area, MRTransformer computes the transformation that maps their behavior to corresponding areas in remote spaces, ensuring consistent interaction dynamics (Figure 1).

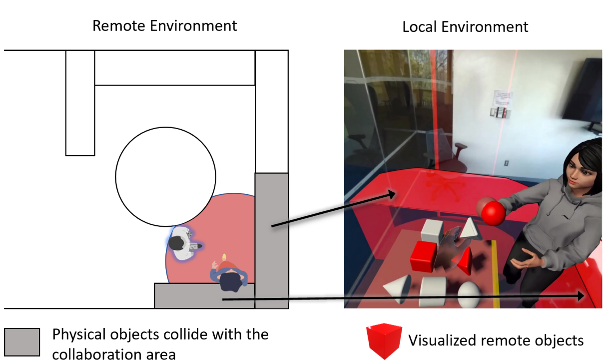

Figure 3: MRTransformer visualizes remote objects as 3D polygons based on overlay computations within collaboration areas, displaying them only when users enter them.

2.2 Avatar Movement Transformations

To transform user movements, MRTransformer applies real-time transformations between collaboration areas in different sites. These transformations are critical for maintaining interaction consistency. When a user repositions a Collaboration Area, MRTransformer recalculates and applies updated transformations, preserving alignment of head and hand movements with remote avatars. The transformations are fed into Final IK's VRIK algorithm to generate full-body avatar movements. Additionally, when users transition between areas, MRTransformer animates avatar movements to provide seamless visual transitions.

2.3 Visualizing Remote Objects

MRTransformer visualizes remote objects as 3D polygons based on overlay computations within collaboration areas, displaying them only when users enter the areas (Figure 3). This approach helps users understand the spatial relationships between real and virtual objects, enhancing spatial awareness and communication efficiency. Using the "Room Setup" functionality, users manually mark physical objects, which are then rendered with transparent red shaders to simulate occlusion.

2.4 Adjusting Collaboration Areas Independently

Users can rotate and move Collaboration Areas independently. For example, rotating an area can resolve avatar occlusions, as the system updates avatar positions to maintain spatial relations (Figure 2). . Moving an area allows users to adjust collaboration zones dynamically. When a user moves an area, the system continuously updates avatar positions, ensuring that the avatars move consistently within the new coordinates, enhancing the flexibility and adaptability of the collaborative environment.

3 User Study

We conducted an exploratory user study to evaluate MRTransformer, comparing it against a baseline system with a global coordinate system. The study included three specific tasks—Avatar Occlusion, Mismatched Location, and Spontaneous Event—highlighting various collaboration scenarios and exploring MRTransformer's challenges and potential.

3.1 Study Environments and Apparatus

The study was conducted in two rooms equipped with shared virtual tables representing collaborative task areas, one for the baseline condition and one for the MRTransformer condition. Each room had its own layout of virtual tables defining collaboration areas and

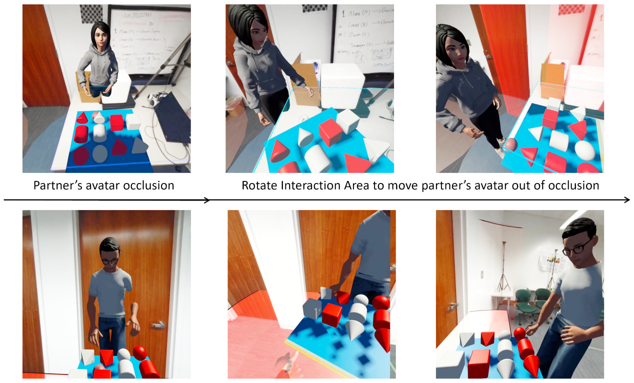

Figure 4: MRTransformer Avatar Occlusion Task — Participants rotate a virtual table to reposition their partner's avatar, using visual cues from 3D red polygons to determine occlusion.

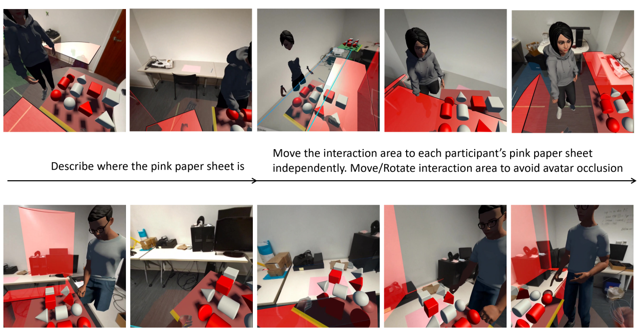

Figure 5: MRTransformer Mismatched Location Task — Participants independently adjust their collaboration areas to their chosen locations

local coordinate systems for MRTransformer, allowing comparison with the baseline condition. Participants wore a Meta Quest Pro (HMD) running our MRTransformer prototype, using hand-tracking for interactions. All devices were connected to the same network, with synchronization and audio via the Normcore library.

3.2 Participants

Twenty participants (10 dyads) were recruited, aged 19 to 30 years ($M = 23.65$, $SD = 2.97$). They had diverse backgrounds and varying levels of experience with VR and AR. Fifteen had used VR headsets before, and three had prior experience with AR/MR headsets. Six participants had experience with platforms like Mozilla Hubs, VRChat, and AltspaceVR.

3.3 Study Tasks

For additional details on the baseline condition, please refer to the supplementary video.

3.3.1 Avatar Occlusion

In this task, participants needed to bring their partner's avatar out of occlusion without moving their collaboration area. In the baseline condition, participants used speech, hand gestures, and a virtual pointer. In the MRTransformer condition, they rotated the virtual table to move their partner's avatar out of occlusion, using visual cues from 3D red polygons (Figure 4).

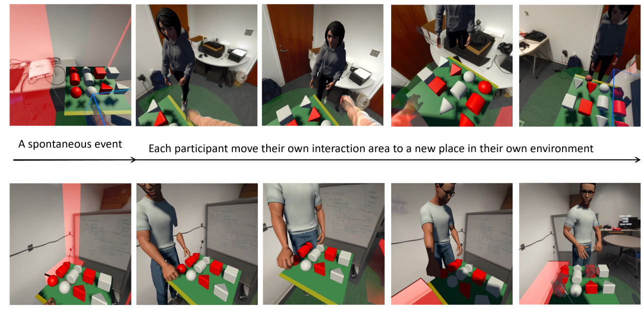

Figure 6: MRTransformer Spontaneous Event Task — Participants independently adjust collaboration areas to prevent avatar occlusion and reposition their partners' avatars as needed.

3.3.2 Mismatched Location

Participants described the location of a pink paper sheet in their room to their partner and moved the shared collaboration area to that location. In the baseline condition, they had to negotiate a single location. In the MRTransformer condition, each participant could independently move the collaboration area to their chosen location (Figure 5).

3.3.3 Spontaneous Event

Participants were asked to move the collaboration area to a new location due to a spontaneous event. In the baseline condition, they moved the table within a shared global system. In the MRTransformer condition, participants independently relocated the collaboration area and adjusted avatars to prevent occlusion (Figure 6).

3.4 Procedure

Participants were guided to separate rooms, given a demo on interaction techniques, and guided through three collaboration tasks. Conditions were counterbalanced using a Latin Square design. After each condition, participants completed questionnaires and semi-structured interviews to discuss their experiences, preferences, and concerns.

3.5 Data Collection and Analysis

Participants completed surveys on social presence and task workload after each condition. The NASA Task Load Index (TLX) was used to assess workload. Semi-structured interviews provided qualitative insights. Thematic analysis was conducted on interview transcripts to generate relevant themes and insights.

4 Findings

4.1 Enhanced Task Efficiency and Satisfaction

Participants preferred MRTransformer for its control and efficiency in preventing avatar occlusion and handling mismatched locations. They reported lower workload scores in MRTransformer compared to Baseline on the NASA TLX, especially in temporal demand, performance, effort, and frustration. As participant P3 noted:

> "...I can just move [partner's avatar] here without telling her where to move like in the previous [Baseline condition]. I guess she can do the same. I don't need to do anything and my avatar can be in the place she wants."

Regarding solving mismatched locations, participants overwhelmingly (85

"You just take few effort and mental workload to do that because when I'm doing the first line like in the first situation [Baseline], I'm thinking about whether the pink paper is physically different in our physical space."

4.2 Transformed Avatar Moments and Social Presence

Some participants preferred the baseline condition for its collaboration feel, though they acknowledged MRTransformer's usefulness. They felt the baseline condition offered a higher social presence, as noted by P11:

"The second one [MRTransformer] is definitely easier because I can have control, but the first one made me feel more like we are in the same space and collaborating. The second one felt like the avatar was controlled by the system."

Others felt MRTransformer maintained social presence even when changing avatar locations, as P5 mentioned:

"[In MRTransformer condition] I really like how her avatar moved with me while I moved the table. It's just like she was actually in my environment and moved with me."

4.3 Concerns about Avatar Movement Transformations

Participants were comfortable with MRTransformer adapting their avatar movements but had concerns about inappropriate uses. Transparency in transformations and informed consent were important. As P16 stated:

I think it depends on the scenario, if it's like today's study I will be totally OK. but I can imagine that if he uses my avatar and creates some funny or inappropriate stuff then I will have concerns.

4.4 Adaptive Visualization for Remote Objects

Participants found visualizing remote objects helpful for understanding their partner's environments and completing tasks. As P2 noted:

Yeah. I really like it. I think it's the most interesting thing I saw today. Seeing it charged in real-time is really cool. I could know where I should stand. I can even know where the table is in his environment. So now I can even put an object on his table.

MRTransformer visualizes remote objects as 3D abstract polygons, visible only if they are inside or partly collide with the collaboration areas. Most participants were comfortable sharing this information. As P17 mentioned:

...I guess it depends on the scenario, if the task requires more details, I don't mind sharing more but if it's not, I prefer something like the red boxes today.

5 Conclusion

In this paper, we introduced MRTransformer, a system designed to enhance remote MR collaboration in incongruent spaces through Dynamic Collaboration Areas. This system allows for real-time adjustments, dynamically aligning avatars' movements with collaborative environments. Our study revealed a strong preference for MRTransformer, especially in handling mismatched locations and adjusting to new interaction scenarios seamlessly. Participants appreciated the enhanced control over the collaboration space, which reduced cognitive load and improved communication efficiency. However, MRTransformer also led to perceptions of reduced personal control over avatar behaviors, potentially affecting social dynamics. These findings highlight the need to balance user perceptions of agency and social presence with improvements in MR remote collaboration experiences.

References

[1] D. I. Fink, J. Zagermann, H. Reiterer, and H.-C. Jetter. Re-locations: Augmenting personal and shared workspaces to support remote collaboration in incongruent spaces. *Proc. ACM Hum.-Comput. Interact.*, 6(ISS), 2022. doi: 10.1145/3567709 1

[2] J. E. S. Grønbæk, K. Pfeuffer, E. Velloso, M. Astrup, M. I. S. Pedersen, M. Kjær, G. Leiva, and H. Gellersen. Partially blended realities: Aligning dissimilar spaces for distributed mixed reality meetings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2023. 1

[3] A. H. Hoppe, F. van de Camp, and R. Stiefelhagen. Shisha: enabling shared perspective with face-to-face collaboration using redirected avatars in virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW3):1–22, 2021. 1

[4] M. Keshavarzi, A. Y. Yang, W. Ko, and L. Caldas. Optimization and manipulation of contextual mutual spaces for multi-user virtual and augmented reality interaction. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 353–362, 2020. doi: 10.1109/VR46266.2020.00055 1

[5] M. Kim and S.-H. Lee. Deictic gesture retargeting for telepresence avatars in dissimilar object and user arrangements. In *The 25th International Conference on 3D Web Technology*, Web3D '20, 2020. doi: 10.1145/3424616.3424693 2

[6] T. Kim, A. Kachhara, and B. MacIntyre. Redirected head gaze to support ar meetings distributed over heterogeneous environments. In *2016 IEEE Virtual Reality (VR)*, pp. 207–208, 2016. doi: 10.1109/VR.2016.7504726 2

[7] N. H. Lehment, D. Merget, and G. Rigoll. Creating automatically aligned consensus realities for ar videoconferencing. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 201–206, 2014. doi: 10.1109/ISMAR.2014.6948428 1

[8] T. Pejsa, J. Kantor, H. Benko, E. Ofek, and A. Wilson. Room2room: Enabling life-size telepresence in a projected augmented reality environment. CSCW '16, p. 1716–1725, 2016. doi: 10.1145/2818048.2819965 1

[9] D. Roth, P. Kullmann, G. Bente, D. Gall, and M. E. Latoschik. Effects of hybrid and synthetic social gaze in avatar-mediated interactions. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 103–108. IEEE, 2018. 1

[10] H. J. Smith and M. Neff. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173863 1

[11] M. Sousa, D. Mendes, R. K. d. Anjos, D. S. Lopes, and J. Jorge. Negative space: Workspace awareness in 3d face-to-face remote collaboration. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*, pp. 1–2, 2019. 1

[12] C. Y. Wang, M. Sakashita, U. Ehsan, J. Li, and A. S. Won. Again, together: Socially reliving virtual reality experiences when separated. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–12, 2020. doi: 10.1145/3313831.3376642 1

[13] J. W. Woodworth, D. Broussard, and C. W. Borst. Redirecting desktop interface input to animate cross-reality avatars. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 843–851, 2022. doi: 10.1109/VR51125.2022.00106 2

[14] B. Yoon, H.-i. Kim, G. A. Lee, M. Billinghurst, and W. Woo. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 547–556, 2019. doi: 10.1109/VR.2019.8797719 1

[15] L. Yoon, D. Yang, J. Kim, C. Chung, and S.-H. Lee. Placement retargeting of virtual avatars to dissimilar indoor environments. *IEEE Transactions on Visualization and Computer Graphics*, 28(3):1619–1633, 2022. doi: 10.1109/TVCG.2020.3018458 1