Transitional Blended Realities: Linking between Video and Mixed Reality for Distributed Team Collaboration

Jens Emil Grønbæk*
School of Computing and Information Systems, University of Melbourne
Department of Computer Science, Aarhus University

ABSTRACT

This paper envisions *Transitional Blended Realities* (TBR)—systems that integrate interfaces for Video Conferencing and Collaborative Mixed Reality. By utilizing a visual overlap in the reference space between remote physical spaces, TBR supports fluid transitions between video conferencing and mixed reality collaboration. This approach has the potential to enhance both collaboration flexibility and accessibility, fostering more dynamic distributed team meetings.

Index Terms: [Human-centered computing Mixed / augmented reality]: Human-centered computing Collaborative and social computing systems and tools—

1 INTRODUCTION

In this era of remote work and global interconnectedness, video conferencing has become an indispensable tool. But traditional video conferencing has limitations when it comes to fostering co-presence among local and remote participants—especially when doing physical tasks such as brainstorming around a whiteboard. In contrast, Mixed Reality (MR) systems offer more immersive and spatially-aware remote collaboration experiences via 3D avatars, but these are yet far from resembling real humans and thus provide limited support natural social interaction. For the foreseeable future, video conferencing will remain relevant, and MR will remain complementary. How do we better integrate the two?

This paper explores a hybrid user interface concept (see Fig. 1) that envisions how we may bridge the gap between video conferencing and MR experiences, allowing for fluid transitions between these for different types of collaboration. We introduce the interaction concept of a visual overlap in the reference space between remote physical spaces as the link for transitioning, creating meaningful associations when users move between video-based and mixed-reality-based collaboration. Two techniques are sketched out to envision how the concept can be realized. The techniques operate across live camera feeds produced by commonplace devices (tablets, desktop computers, etc.) and mixed reality headsets. They are based on the idea that users can create meaningful links between remote spaces by overlapping videos of themselves and their local task spaces to produce transitional remote collaboration spaces.

2 BACKGROUND

The concept presented in this paper comes from my background in developing new flexible interfaces for synchronous distributed team collaboration. A key challenge is to support users in effectively configuring collaborative interfaces for work while participants are distributed across dissimilar spaces. To address this challenge, I have been investigating the concept of supporting end-user configurability: How can the spatial arrangement of collaborative user interfaces be more configurable by the end-users? The research direction is inspired by studying socio-spatial dynamics in various co-located and distributed settings. In co-located meetings, participants often configure physical meeting rooms by rearranging the furniture and display technology to accommodate different meeting situations. In my research, I explore how interfaces for remote and hybrid meeting settings can be designed to support participants in "rearranging" the meeeting space of video conferences and mixed reality technologies.

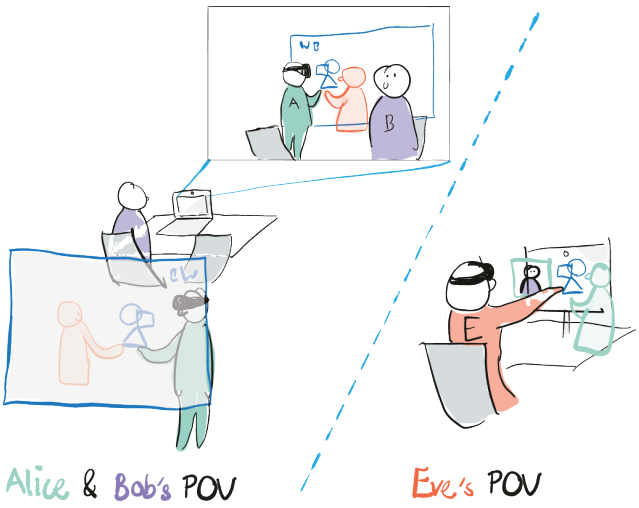

Figure 1: Using Transitional Blended Realities (TBR), Alice, Bob, and Eve can switch between video and mixed-reality interfaces through links in the blended space between different physical task spaces.

First, our recent work on MirrorBlender [3] (see Fig. 2, left), along with a range of recent commercial systems (Gather, Sprout, Wonder, Teamflow, etc.), represent a new paradigm in video-conferencing interfaces, where users control an avatar-like representation of themselves to feel more embodied in the shared virtual space. What is unique about MirrorBlender is the users' ability to not only move their mirrors (videos) around in the shared workspace, but also overlap these into a shared blended interaction space.

Second, while collaborative MR fosters a sense of physical co-presence not possible in video conferencing, working across different spaces causes alignment issues due to physical layout dissimilarities. In our recent paper [4], we propose that users can configure Partially Blended Realities, i.e., blended spaces that incorporate their local physical surfaces as shared virtual surfaces (see Fig. 2, right). Key here is the flexibility that is derived from only blending the *relevant parts* of their respective spaces, such as their local desks and whiteboards.

3 VIDEO AND MIXED REALITY ARE COMPLEMENTARY

Although the interface concepts of MirrorBlender and RealityBlender can adapt to a range of collaborative situations, video and MR are fundamentally different and complementary media. To understand their complementary nature, we may turn to Buxton's distinction of *person, task, and reference spaces* [2]. Video conferencing puts emphasis on person space with the ability to read other's facial expressions, while mixed reality enables better physical task

*e-mail: jensemil@cs.au.dk

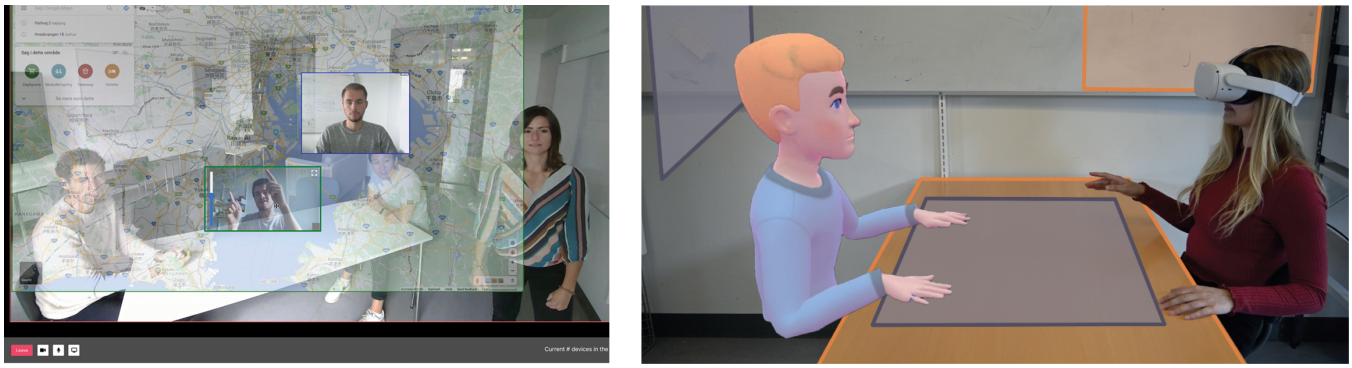

Figure 2: Left: MirrorBlender (reprinted from [3]). Right: RealityBlender (reprinted from [4])

space sharing with the ability to anchor virtual content meaningfully across distributed spaces. The reference space, then, is the intersection between person and task space, and video as well as MR tend to compromise either the person or task space experience. As users switch between person-focused and task-focused conversations, the collaborative interfaces should provide ways for users to seamlessly transition between interface modes that best suit the task at hand.

MirrorBlender and RealityBlender make up important components in providing an accessible and flexible hybrid user interface for distributed collaboration. But they have yet to be integrated in a coherent hybrid interface. The rest of this paper will explore this idea, building on the notion of transitional interfaces [1]—systems that let users switch between interface types, depending on the nature of the collaborative task. A recent study of such interfaces by Schroeder et al. [8] led to a comprehensive taxonomy of patterns on how collaborating dyads switch between interface types. While this research shows collaborative benefits of transitional interfaces, there are still many open questions about how to design hybrid user interfaces for distributed collaboration spaces.

4 Transitional Blended Realities

This paper introduces Transitional Blended Realities (TBR). In TBR, the idea is to integrate MirrorBlender- and RealityBlender-type interfaces into a coherent system that would enable participants in distributed team meetings to switch effortlessly between these as two modes. Albeit different (one is 2D video, the other 3D mixed reality), these two interfaces share the same fundamental philosophy—that collaborating users benefit from the ability to actively participate in the configuration work of blending their distributed spaces as they shift between different modes of collaboration. In the forthcoming, let us refer to such configuration work as either happening through video (MirrorBlending Mode) or MR (RealityBlending Mode).

TBR systems incorporate the following three interface principles:

- **Multiple Avatar Representations:** Participants are represented by their respective avatars (either floating video windows or virtual 3D embodied avatars), allowing seamless continuity when transitioning between the two modes.
- **Blended Reference Space:** The hybrid user interface establishes a blended reference space accessible to participants both in MirrorBlending and RealityBlending modes. In this space, both 3D avatars and video mirrors can coexist and interact.
- **Transition Portal:** To facilitate transitions, TBR provides a transition portal within the shared reference space. This portal serves as a link between the two modes. When a participant in MirrorBlending Mode desires to switch to RealityBlending Mode, they perform a gesture upon wearing the headset. The interface transitions then differ depending on whether a given user is co-located or remote with respect to the user initiating the transition, and which mode the user is currently in.

5 User Scenario

To demonstrate the TBR interface principles in action, we describe a user scenario of distributed team collaboration between Alice, Bob, and Eve using TBR for a brainstorming task:

Setting Up the TBR Meeting: Alice and Bob are co-located team members working on a brainstorming task in their office. They have started jotting down ideas on the local whiteboard. Recognizing the value of Eve's expertise and creativity, who is currently remote, they decide to involve her using TBR. Standing in front of the whiteboard, they initiate the TBR distributed team meeting in MirrorBlending Mode to establish a shared reference space around the whiteboard.

MirrorBlending Mode with Remote Expertise: Eve joins remotely using the TBR MirrorBlending Mode. Her mirror (live video feed) appears within the shared reference space, blended seamlessly with the feed of Alice and Bob (see Fig. 3A).

Distributed Brainstorming Session: Alice and Bob share their initial ideas on the local whiteboard, discussing and refining them. They involve Eve in the conversation. Through the TBR MirrorBlending Mode, she can see how they gesture around the whiteboard and actively participate in the brainstorming session by pointing with hear hands through her projected mirror image.

Alice Shifts to RealityBlending Mode: As the brainstorming session progresses, Eve becomes more active, and Alice finds it inconvenient to look at Eve and the whiteboard through the laptop screen. She puts on a headset and transitions to RealityBlending Mode by directly dragging the mirror images from the screen onto the real whiteboard in her own view (see Fig. 4A-C). In the RealityBlending Mode, Alice can now perceive Eve's pointing gestures correctly located at the real whiteboard.

Eve Shifts to RealityBlending Mode: Eve starts to find pointing through the mirror image insufficient for contributing to the ideation. She puts on the headset which detects her own mirror location on the whiteboard (see Fig. 3B). Bob, who is still following along via MirrorBlending Mode, perceives Eve as transforming from a mirror image into a 3D avatar (see Fig. 3C). Eve's screen becomes a local placeholder for the remote physical whiteboard, where Bob is recreated as a cropped mirror image and Alice appears as an embodied avatar (see Fig. 3D).

Returning to MirrorBlending Mode: Finally, Alice and Eve take off the headsets and reengage in person space communication with Bob (in the MirrorBlending Mode) to wrap up the meeting.

6 Discussion and Open Questions

The above user scenario focuses on demonstrating the value in having the flexibility to transition between interface modes to accomodate for the dynamic physical nature of collaborative work.

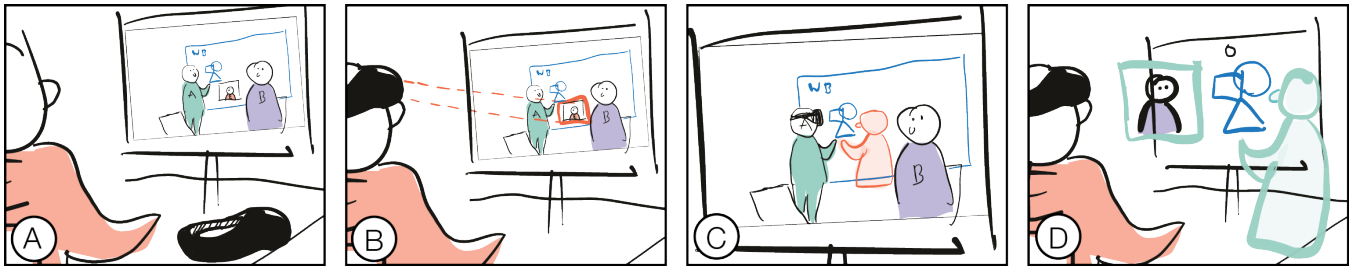

Figure 3: *Cross-device interaction from Eve's POV.* A: Eve fails to point effectively through her video mirror image. B: She switches to MR and blends her reality with theirs. Her headset detects the location of her mirror (video window) on the screen. C: Her avatar then appears in this location in both Alice's and Bob's POV (shown in Fig. 1). D: Eve now sees Alice as a 3D avatar while maintaining a view of Bob's mirror image.

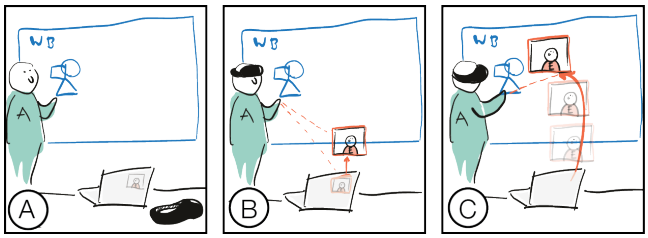

Figure 4: *Cross-device interaction from Alice's POV.* A: Alice has switched to her physical whiteboard and wants to bring Eve into her physical task space. B: She pulls Eve's mirror out of the screen to drag and drop it at the whiteboard. C: Eve's mirror gets anchored next to the relevant diagram sketch.

The interactions shown in Fig. 3 and Fig. 4 combine a headset and an external display in the meeting space, exploring the metaphor that users can use direct manipulation of avatar representations to configure appearance of meeting participants—initated either by configuring one's own representation in the remote space (Fig. 3) or a collaborator's representation in one's local space (Fig. 4). While these examples sketch out the possibilities, there are still many details of these interactions that have yet to be fleshed out, many of which would be discovered when engaging with the concept through prototyping. However, they help open up discussion of the following important questions and research challenges.

- **Technical/infrastructural challenge:** *Which ecosystem requirements are needed?* What sensing and detection capabilities are feasible to require for different devices to enable the transition portals proposed in the above techniques? How would the techniques work across a wider set of devices and sensing capabilities (e.g., MR phones and meeting displays)?

- **Interaction design challenge:** *Transition flexibility increases system complexity.* How do we reduce the effort of configuration work without reducing flexibility? Moreover, how do we design transitions from different users' POVs (video vs. MR) without breaking consistency in the group's mental model of the shared blended space?

- **User experience challenge:** *How can such hybrid user interfaces enable remote users to feel more included in distributed meetings?* Primay room dominance is a key challenge in hybrid meetings [7]. Can we make remote users feel more included by giving them the flexibility to switch devices for different kinds of embodiment in the primary space? And how would we evaluate the success of such a user interface?

- **Conceptual challenge:** *Which theories do we turn to to inform design choices and understanding collaborative behavior in such hybrid environments?* A unifying theory here could be proxemics [5], which studies the role of physical space (such as furniture elements) in social interaction. It may help inform how to develop techniques for anchoring virtual avatars [4] and shared virtual objects [6], and how hybrid user interfaces may incorporate Space-Function-Distance relationships [2].

7 Conclusion

The integration of video and mixed reality into a hybrid user interface holds great promise for transforming distributed team meetings. We have explored the notion of Transitional Blended Realities by sketching out interaction concepts and scenarios where distributed collaborators use transition portals in their blended space to switch between modes of video conferencing and mixed reality. Supporting these transitions can enable users to combine the complementary benefits of both modes for more dynamic distributed meetings.

References

[1] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook: a transitional ar interface. *Computers & Graphics*, 25(5):745–753, 2001. Mixed realities - beyond conventions. doi: 10.1016/S0097-8493(01)00117-0

[2] B. Buxton. *Mediaspace – Meaningspace – Meetingspace*, pp. 217–231. Springer London, London, 2009. doi: 10.1007/978-1-84882-483-6_13

[3] J. E. Grønbæk, B. Saatçi, C. F. Griggio, and C. N. Klokmose. Mirrorblender: Supporting hybrid meetings with a malleable video-conferencing system. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445698

[4] J. E. Grønbæk, K. Pfeuffer, E. Velloso, M. Astrup, M. 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*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581515

[5] E. T. Hall. *The Hidden Dimension*. Doubleday, New York, NY, USA, 1966.

[6] W. Luo, A. Lehmann, H. Widengren, and R. Dachselt. Where should we put it? layout and placement strategies of documents in augmented reality for collaborative sensemaking. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3501946

[7] B. Saatçi, K. Akyüz, S. Rintel, and C. N. Klokmose. (re)configuring hybrid meetings: Moving from user-centered design to meeting-centered design. *Computer Supported Cooperative Work (CSCW)*, 29:769–794, 2020. doi: 10.1007/s10606-020-09385-x

[8] J.-H. Schröder, D. Schacht, N. Peper, A. M. Hamurculu, and H.-C. Jetter. Collaborating across realities: Analytical lenses for understanding dyadic collaboration in transitional interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3580879