

Figure 1: Examples of different Social MR/AR/VR experiences that strove to convey a real-time telepresent capture of participants, from group experiences in CAVE environments [1], to projection-mapped spatial AR [9], to fully realized 3D “holoportation” [3], to telepresence in VR [7].

Shared, Synchronous, Social Mixed Reality Experiences

Mark McGill

School of Computing Science
University of Glasgow
mark.mcgill@glasgow.ac.uk

Stephen Brewster

School of Computing Science
University of Glasgow
stephen.brewster@glasgow.ac.uk

Florian Mathis

School of Computing Science
University of Glasgow
florian.mathis@glasgow.ac.uk

Abstract

Our position for the Social VR workshop is that the remit should be expanded to more broadly consider Mixed Reality (MR) – (a)synchronous communication at-a-distance is not exclusively limited to visually-oriented telepresence delivered through VR HMDs. Rather, there is a space within which we might facilitate shared MR-driven experiences visually (using traditional AR/VR/MR headsets) and aurally (e.g. using auditory headsets), synchronously (same time) and asynchronously (different times), in the same place or at-a-distance, and asymmetrically (e.g. with mixed headset types), with a variety of permutations of these factors – and perhaps the most impactful permutations may not be grounded in VR headset-driven experiences.

Author Keywords

Mixed Reality; Social; Telepresence; Virtual Reality; At-a-distance;

CCS Concepts

•Information systems → Collaborative and social computing systems and tools; •Human-centered computing → Virtual reality;

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).
CHI'20, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
<https://doi.org/10.1145/3334480.XXXXXXX>



Figure 2: Examples of more VR-oriented social experiences where avatars are used to convey social presence, giving users more control over their appearance, but conveying less innate social cues. From top: *Oculus Social Alpha*^a from 2015; *VR Chat*^b; *Facebook Horizon*^c; and Facebook's *Codec Avatars: Conversation in VR*.

^aengadget.com/2015/10/28/oculus-social-alpha-delivers-group-attachng-to-virtual-reality/

^bvrchat.com/

^coculus.com/facebookhorizon

The Social Story So Far

Our capability to communicate and share experiences at-a-distance has changed markedly over the last decade. Regarding communication, our homes are now filled with directional microphones attached to smarthome devices, whilst our smartphones, tablets, MR headsets and TVs have all featured bi-directional audio/visual communication capability for some time. Our capability to engage in shared, synchronous experiences whilst using these communication modalities has also changed significantly - from video conferencing with screen share capabilities, to immersive virtual experiences in gaming, simulation, education and training where all can be “telepresent” regardless of location. However, whilst social VR has seen significant growth (e.g. the meteoric rise of apps like VR Chat, and Facebook’s equivalent moves with Spaces/Horizon), despite its significant benefits (e.g. in terms of facilitating embodied telepresence and the sensation that the person is actually present, or conveying avatars that can relay speech and a subset of social cues through head/hand movements in a 3D space) social VR has yet to make a significant breakthrough in altering how the public at-large communicate in their day-to-day lives. Why is that, and why do we argue that the remit of the workshop should broaden toward social MR more generally?

Social VR, AR, MR/XR?

Social VR encompasses a broad set of experiences. There is no one defacto standard for what should constitute a social VR experience, how it is enacted, how distant others are represented/conveyed etc. Social VR is an object of fixation for both academia and industry because the kinds of presence that VR headsets can facilitate: presence in the immersive, place/plausibility illusion sense [10]; and social presence i.e. the perception that the other persons present are real and alive. Thus VR as a medium can allow us to

escape our physical surroundings, and indeed our physical form, inhabiting a new form of our choosing/design that will be our virtual proxy through which our interactions with others are mediated. This is a powerful, potentially intoxicating combination, given the previous impact of social gaming experiences (e.g. MMORPGs) in previous decades - enabling a more perceptually real, more immersive social experience at-a-distance. However, VR is also undeniably restrictive [6], being occlusive of our surrounding real-world environments and collocated others. There are other delivery modalities for social experiences across the mixed reality continuum which may often be better suited to the facilitation of social connections and experiences, be they (a)symmetric, (a)synchronous, or co-located/at-a-distance, or any permutation/intermixing of these attributes.

If we assume that the “social” element refers to a conveyance of the *{presence, social signals, intents, actions, messages and other related constructs}* of others that are separated from the perception of the user, this could encompass everything from a VR headset user cut-off from those proximate others in the same space at the same time, to an AR headset user communicating with others located in a different space, even at an altogether different time, and everything in-between. What is communicated might vary from the more abstract (a partner leaving a spatial, geo-located message triggered on your walk home) to variations in spatialized auditory presence [8] and avatar-based presence, all the way to fully captured and embodied real-time telepresence (e.g. Holoportation [3] or our own work [7]). Given this, it would seem churlish to consider VR alone. The internet and the smartphone both revolutionized our interpersonal communications, giving us the flexibility and freedom to switch modalities near-instantly from text, to voice, to video, a freedom that different societies/cultures have taken to very differently (e.g. the use of audio mes-



Figure 3: Examples of shared at-a-distance VR experiences: *Convrge Cinema*^a where users could watch films together at-a-distance; *UIST 2019 VR poster session*^b where virtual attendees could take part in a purely VR poster session at UIST; and *BigScreen VR*^c where users could selectively share, and collaborate using, captured OS desktop instances/applications.

^a<https://www.roadtovr.com/convrge-launches-social-home-theatre-space-convrge-cinemas/>

^btwitter.com/florian_mathis/status/1186764399563935744

^cbigscreenvr.com/

saging in some countries over text messaging). Consequently, any changes in the *de facto* personal computing experience are highly likely to in-turn feed into changes into how we interact with others at-a-distance as well.

In contrast to VR, Mixed/Augmented Reality systems will have the capacity to seamlessly integrate virtual content into both reality, and our everyday lives, through wearable (nay fashionable) headsets designed and intended to be worn continuously. What this means for how we interact with others is not yet understood, but we would suggest this is a key topic of interest. Our own research at CHI this year looks at auditory MR specifically [8], and we could envisage a permanent auditory feedback modality allowing for a variety of new and altered ways of communicating e.g. spatialized speech at-a-distance for a sense of presence when talking to distant others; changing attitudes toward text messages given auditory delivery (perhaps even in the sender’s simulated voice); changing attitudes towards voice chat (given the popularity of Discord for example - might you have a “drop-in” room that you enter when commuting to talk to friends day-to-day?); and even the ability to leave spatialized, location-based auditory messages. If more visually-oriented AR can catch up with consumer expectations (already amplified by the comparable capability of affordable consumer VR headsets), what sea change’s might we see precipitated by such technology, or rather – what sea changes can we envisage and design for that might encourage adoption of said technology? Consequently, we would argue that the scope of the workshop should, in-part, try to consider a broader remit of social MR more generally - of which VR is only one important facet.

The Demands of Social MR

In relation to our previous work, we also see significant efforts in facilitating social MR experiences:

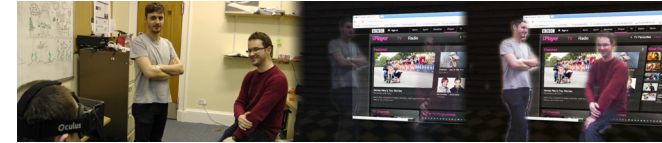


Figure 4: Our 2015 work [6] bringing proximate persons into VR.

Designing for Different Social Activities

If an activity is shared, the activity too places certain demands and constraints regarding how communication is facilitated. For example, in our previous work we used consumer RGB depth cameras to visualize collocated persons [6] and at-a-distance others in shared, social VR co-viewing experiences [7] (e.g. watching a movie in a cinema together, watching a 360 video together, see Figure 5). Arguably, with the shared focal point of a cinema screen (with the social norms that such an environment suggests), the necessity of tracking and conveying the whole participants body is debatable, compared to avatar-based presentation, as users would perhaps be unlikely to wildly gesticulate, and are in a fixed, seated position. However, co-viewing a 360 degree video, whilst a very similar task, offers much greater opportunity for expression during the experience (e.g. see the person recoiling from a virtual shark in Figure 5), and consequently demands a greater capacity to convey the social signals. Indeed, such differences might be exhibited across activity domains. For productivity, basic avatars that convey the focus of attention might be sufficient for facilitating a typical conference call - but for designers interacting with the latest virtual model of their product, the demand is on the VR/MR technology to approximate the experience of our distant others actually “being there” - from gesticulations, to facial expressions, to the focus of their gaze, to even the intended direction of their utterances.

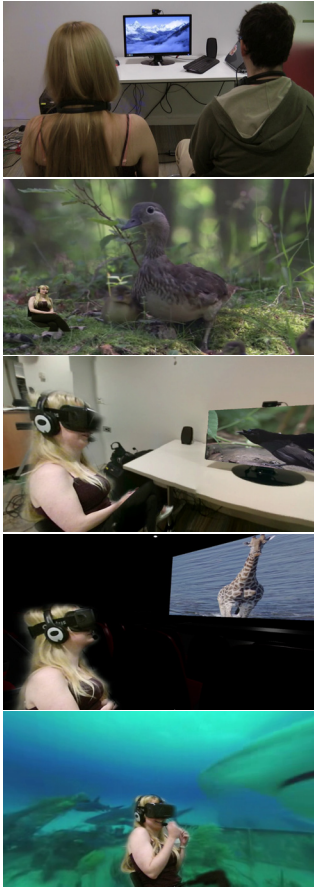


Figure 5: Images from our shared at-a-distance media experiences paper [7], where users experienced different forms of media together with different at-a-distance presentations, including TV-based PiP and VR-based telepresence in real-time captured via Kinect V2 cameras.

Public, Semi-Public and Private Collocated Interactions

Nor is it enough in such scenarios to facilitate presence alone, with the conveyance of, and interaction with, virtual content being a necessity. There are challenges here from the logistical (e.g. how best to share application instances/views and other media/content across users in a scalable, low-latency way) to those of space (e.g. how best to facilitate multiple viewing of content; how to translate and scale virtual experiences between different physical spaces) to those of interaction with virtual content (e.g. creating private duplications and instances for transient interactions) and control over the visibility of interactions with others (e.g. do I want others in the room to know I am not attending to the presentation when I can hide that fact? Can I have private conversations without having to whisper? Can I take entirely private notes?).

All of these issues pose significant challenges to computing and HCI more generally, lest the potential of such social MR experiences be lost due to the perceived complexity of engaging in these experiences effectively. But these challenges also suggest that social MR experiences have the potential to go beyond what is possible in physically co-present interactions (e.g. private whispering that cannot be overheard). HCI can lead the way in re-envisioning how we communicate and interact through these new modalities.

Social Acceptability and Security of Sensing

Neither can we ignore the physical hardware requirements to create experiences at the far end of this spectrum – any embodied telepresence experiences first require a digital twin of the user to be conveyed (be it an avatar, or a prior/real-time captured likeness e.g. through 3D body/face scan¹ or imaging, or some combination). Then, sensing is required to capture the relevant information about us and

¹techcrunch.com/2018/05/02/facebook-photo-realistic-avatars/

our actions from the environment - this might be head position/orientation through the headset, to capturing the full body pose and facial expressions, to selectively capturing parts of our visage (e.g. fingers/hands, gaze), to capturing the objects we interact with, to capturing hidden/unseen information (e.g. biometrics conveyed to the doctor or running partner at-a-distance). And this must be accomplished across all the co-located users intended to be conveyed (e.g. the family in the living room talking to a distant relative). This is a challenge not just in terms of cost or logistics – but in terms of the social acceptability and security of integrating and distributing this kind of sensing into our environments. How can users be safeguarded, such that they can engage in the best at-a-distance experiences possible without compromising their day-to-day privacy and security? And will users ever be persuaded that the measures taken to safeguard them will be enough (e.g. the backlash against smart speakers despite Amazon and Google both employing hardware wakeword detection)? For example, what abusive deep-fakery might be enabled if your personal telepresence 3D facial scan was acquired by others? What additional threats on users' security, privacy, and safety (similar to [5]) emerge when participating in collocated social MR experiences in public or semi-public spaces?

Our Position

There is a gamut of MR technology that could better facilitate communication and shared experiences at-a-distance. We would argue that rather than focussing on VR (highly relevant for consumers right now), we need a framework for better describing the social MR space, and the new grand challenges/potentials enabled therein e.g. attitudes to the conveyance of others, designing shared interactions and experiences, sensing requirements, ethics, security and privacy, and consideration of intermixed (a)symmetric (e.g. [4]) and (a)synchronous (e.g. [2]) social experiences.

REFERENCES

- [1] Stephan Beck, André Kunert, Alexander Kulik, and Bernd Froehlich. 2013. Immersive group-to-group telepresence. *IEEE transactions on visualization and computer graphics* 19, 4 (apr 2013), 616–25. DOI : <http://dx.doi.org/10.1109/TVCG.2013.33>
- [2] Pablo Cesar, Konstantinos Chorianopoulos, and Jens F. Jensen. 2008. Social television and user interaction. In *Computers in Entertainment*, Vol. 6. 1. DOI : <http://dx.doi.org/10.1145/1350843.1350847>
- [3] Sean Fanello, Sergio Orts-escolano Christoph Rhemann, Mingsong Dou, Vladimir Tankovich, Charles Loop, and Philip Chou. 2016. Holoportation : Virtual 3D Teleportation in Real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. 741–754. DOI : <http://dx.doi.org/10.1145/2984511.2984517>
- [4] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 4021–4033. DOI : <http://dx.doi.org/10.1145/3025453.3025683>
- [5] Florian Mathis and Mohamed Khamis. 2019. Privacy, Security and Safety Concerns of using HMDs in Public and Semi-Public Spaces. In *Proceedings of the 1st Workshop on Challenges Using Head-Mounted Displays in Shared and Social Spaces at CHI 2019*. Association for Computing Machinery, New York, NY, USA, 5.
- [6] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proc. of CHI '15*. ACM Press, New York, New York, USA. DOI : <http://dx.doi.org/10.1145/2702123.2702382>
- [7] Mark McGill, John H. Williamson, and Stephen Brewster. 2016. Examining The Role of Smart TVs and VR HMDs in Synchronous At-a-Distance Media Consumption. *ACM Trans. Comput.-Hum. Interact.* 23, 5, Article Article 33 (Nov. 2016), 57 pages. DOI : <http://dx.doi.org/10.1145/2983530>
- [8] Mark McGill, Graham Wilson, David McGookin, and Stephen Brewster. 2020. Acoustic Transparency and the Changing Soundscape of Auditory Mixed Reality. In *To Appear In Proceedings of ACM CHI 2020*.
- [9] Tomislav Pejisa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew D Wilson. 2016. Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*. ACM Press, New York, New York, USA, 1714–1723. DOI : <http://dx.doi.org/10.1145/2818048.2819965>
- [10] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364, 1535 (dec 2009), 3549–57. DOI : <http://dx.doi.org/10.1098/rstb.2009.0138>