

A Mixed Reality Approach for dealing with the Video Fatigue of Online Meetings

Vigneswaran Poobalasingam^{1,2}, Julie Wall¹, Saeed Sharif¹, Mansour Moniri¹, Cornelius²
Glackin, Nigel Cannings²

¹University of East London, UK

²Intelligent Voice, UK

Abstract

Much of the issue with video meetings is the lack of naturalistic cues, together with the feeling of being observed all the time. Video calls take away most body language cues, but because the person is still visible, your brain still tries to compute that non-verbal language. It means that you're working harder, trying to achieve the impossible. This impacts data retention and can lead to participants feeling unnecessarily tired. This project aims to transform the way online meetings happen, by turning off the camera and simplifying the information that our brains need to compute, thus preventing 'Zoom fatigue'. The immersive solution we are developing, iVXR, consists of cutting-edge augmented reality technology, natural language processing, speech to text technologies and sub-real-time hardware acceleration using high performance computing.

Keywords: Augmented Reality, Virtual Reality, Online Meetings, Mixed Reality

Company Description

The company is a global leader in the development of proactive compliance and eDiscovery technology solutions for voice, video, mixed-reality and other media. The core business of the system is speech recognition and natural language understanding technology, providing complex analytic capabilities of speech audio. Its clients include government agencies, banks, securities firms, Call-Centers, litigation support providers, international consultancy, advisory businesses and insurers, all involved in the management of risk and meeting of multi-jurisdictional regulation.

Fundamental to its success, its patent-pending and patented technologies are developed by a team of dedicated researchers and system engineers based in the UK. The company leads the market and maintains its strengths in the areas of thought leadership, innovation, R&D and providing solutions to its clients.

Project Summary

This project aims to transform the way online meetings happen with cutting-edge augmented reality technology, natural language processing, speech to text technologies and sub-real-time hardware acceleration using high performance computing.

Much of the issue with video meetings is the lack of naturalistic cues, together with the feeling of being observed all the time. During in-person communication, our brains use the 7-38-55 rule to decipher the meaning behind what's being said. 7% verbal, 38% tone of voice and 55% body language [1]. Video calls take away most body language cues, but because the person is still visible your brain still tries to compute that non-verbal language. It means that you're working harder, trying to achieve the impossible. This impacts data retention and can lead to participants feeling unnecessarily tired.

At the least, part of the answer could lie in the growing world of VR gaming. In VR gaming, the action is driven by the user's actions, usually through a gamepad or keyboard. The visual representation of the person is an avatar. This allows the human operator to have a degree of distance from their online presence. So, if they need to scratch their nose, the avatar is not mimicking them. On a video call, all actions are immediately transmitted and seen by other participants and that puts extra pressure on everyone involved. But of course, meetings are not about shooting and fighting. This means that the way we interact will be very different and will need a different type of control, this is where the voice becomes so important.

To avoid unnecessary user actions, we want to be able to turn off the camera completely. We need the actions of the avatars to be driven by what the user is saying. The speech recognition and speech analysis expertise of the project partners allow us to understand what is being said, and how it is being said. If the tone is light and friendly, the avatar will relax and smile. If the tone is aggressive, it may lean forwards to make a point. We can seat everyone naturally round a boardroom table, or we can all sit on the deck of a yacht to sip cocktails. The setting is irrelevant. The point is that it simplifies the information that our brains need to compute, preventing 'Zoom fatigue' [2]. Additionally, moving away from standard video conferencing, reduces the bandwidth requirements of the application.

Project Details

The technology we are developing, called iVXR, uses both the Android and iOS developer kits to build an immersive meeting experience [3]. The product can support mono single screen mode for a hand-held AR/VR meeting, mixed-reality (MR) mode using Oculus Quest 2 and supports the inexpensive cardboard options with split screen stereo AR or VR options. We are using the Unity game engine as the basis of the technology, a proven solution for massively multi-user applications, and this together with the inexpensive MR technology we prefer is intended to provide a useful application for a wider community of potential users. The user's avatar can be configured and personalised using an authenticated UI. The application offers fine-grained personalisation of the avatars, around appearance, gender, ethnicity and clothing. The avatars are animated in real time, their facial expression system is based on the Facial-Action-Coding-System (FACS) [4].

iVXR supports four different application modes, 3D Mono, AR Mono, 3D Stereo and MR Stereo. 3D Mono is a standard 3D rendering in a standard display, running on Windows, Android, macOS, and iOS. For AR Mono, the CGI image is augmented onto the actual world image and rendered in a standard display. AR Mono is supported on mobile devices that run Android with ARCore support or iOS with ARKit. 3D Stereo, also known as VR, renders two different visuals from the viewing points of the eyes, currently supported on the Oculus Quest 2. Finally, MR Stereo, also known as 'Passthrough' mode on the Oculus, is an experimental feature that shows the natural world as a grayscale background, with the CGI overlaid on top. This mode renders two different visuals like VR. Participants can choose any of these application modes to participate in the same meeting.

The application provides near real-time subtitling of the speech audio using the company's automatic speech recognition engine and natural language processing technology. Speaker-separated transcripts can be emailed, copied or saved after the meeting, serving as a permanent record.

Feedback from End Users

Regular user testing sessions have taken place amongst the project partner's staff to elicit feedback on and to support iterative refinement of the application. In terms of usability and user experience, the workshops focused on:

- Ability of the users to join the meeting from the various supported platforms

- Ease of personalisation of the avatar and configuration of the virtual environment
- Technical feedback on latency, responsiveness, sound quality and transcription accuracy
- Naturalness of the shared experience

Our users have expressed high levels of satisfaction, while identification of practical issues around usability have informed ongoing development, while feedback on user experience has indicated positive responses to feelings of engagement with iVXR’s virtual environment. Increasingly more productive meetings amongst the project partners using iVXR are taking place, reflecting the improvement through this ongoing consultation process.



Figure 1: AR Mono application mode, where the CGI image is augmented onto the actual world image and rendered in a standard display

Future Outlook

Attempts have been made during the pandemic to try to change the current 2D imagery of video meetings and make it more accessible. Teams “Together” mode is an example. But these endeavours don’t resolve the underlying problems of “presence”. There are numerous competing solutions in various stages of development in the market place, such as Spatial [5], Horizon Workrooms [6], MeetinVR [7], Glue [8], Mozilla Hubs [9], BigScreen [10], ENGAGE [11], Rumii [12], AltspaceVR [13], Rec Room [14], and FrameVR [15]. However, it is difficult to separate the conceptual from the real implementations. The ultimate aim is to develop a security-conscious technology that integrates with the daily business workflow.

In this project, we have focused on providing an immersive meeting experience, with reduced bandwidth by concentrating on the audio channel, cross-platform support, built in speech-recognition and natural language processing technology. Our immediate future focus is to provide secure communication using this application, by incorporating the company’s patented privacy preserving technology for audio communication.

Conclusion

Technologies like the Oculus provide a vehicle for this application to reach a wider audience. As more AR applications become available for the end user, the hardware to support these will continue to reduce in price and become increasingly accessible. Having started this project pre-pandemic, the application described here is timely and provides a viable alternative approach to standard video-conferencing, promoting more effective and efficient meetings.

References

1. B. Library, "Albert Mehrabian," British Library, <https://www.bl.uk/people/albert-mehrabian>
2. Fosslien, L. and Duffy, M.W., 2020. How to combat zoom fatigue. Harvard Business Review, 29.
3. Abdullah Ali, Cornelius Glackin, Nigel Cannings, Julie Wall, Saeed Sharif, Mansour Moniri, "A Framework for Augmented Reality Based Shared Experiences", Immersive Learning Research Network Conference (iLRN), 2019.
4. B. Farnsworth, "Facial Action Coding System (FACS) – A Visual Guidebook," <https://imotions.com/blog/facial-action-coding-system/>
5. Spatial Systems, <https://spatial.io/>
6. Facebook Technologies, LLC., "Meta Quest Workrooms," <https://www.oculus.com/workrooms/>
7. MeetinVR, <https://www.meetinvr.com/>
8. Glue, <https://glue.work/>
9. Hubs, Mozilla, <https://hubs.mozilla.com/>
10. bigscreen, <https://www.bigscreenvr.com/>
11. Engage, <https://engagevr.io/>
12. Doghead Simulations, <https://www.dogheadsimulations.com/>
13. Microsoft inc, AltspaceVR, <https://altvr.com/>
14. Rec Room, <https://rec.net/>
15. Frame, [Online]. Available: <https://framevr.io/>