Projecting "The Extended Mind:" Expanding Visual and Interactive Space for Remote Collaboration Experiences in Mobile Mixed Reality

Bektur Ryskeldiev
University of Tsukuba
Japan
bektour@slis.tsukuba.ac.jp

Yoichi Ochiai
University of Tsukuba
Japan
wizard@slis.tsukuba.ac.jp

ABSTRACT

This study investigates how pervasive mobile mixed reality telepresence can assist users in collaborative tasks, entertainment, and navigation, by expanding their visible and interactive space within a shared mixed reality environment. To address this issue, this work proposes and discusses the Spatial Livestream Composition (SLC) method that combines both directionalized live media streams and photospherical imagery in a single collaborative multiuser environment. The proposed method was implemented in three proof-of-concept applications, and two of them were validated in user studies. The preliminary results have shown decrease in mental workload, increase in spatial and situational awareness, as well as increase in user engagement in comparison with similar applications with regular, non mixed reality, interfaces. The rest of this paper discusses current and future applications of proposed method.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**; *User interface management systems*;

KEYWORDS

Spatial Media, Mixed Reality, Social Media, Telepresence, Mobile Computing, Groupware, Crowdsourcing, Photospherical Imagery

ACM Reference Format:

Bektur Ryskeldiev and Yoichi Ochiai. 1997. Projecting "The Extended Mind:" Expanding Visual and Interactive Space for Remote Collaboration Experiences in Mobile Mixed Reality. In *Proceedings of ACM Woodstock conference (WOODSTOCK'97).* ACM, New York, NY, USA, Article 4, 3 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

In "The Extended Mind," Clark and Chalmers claim that humans use external objects and environments to aid and extend their cognitive functions [3]. Similarly, in the current age of social media, live streaming, and mixed reality interactions, one might argue that we are now adopting real and virtual remote objects and environments as tools that assist our cognition, projecting the extended minds even further outside of our immediate physical surroundings.

Such tools, however, are still rather primitive. For instance, most social and collaborative applications still use single-viewpoint, rectangular live video streaming as the main form of interaction, which limits users' visible and interactive space (Figure 1). Therefore, this study aims to answer on the following questions: *"How can we extend the visible and interactive space available to users?"*, and *"How can we take advantage of extended interactive space in social and collaborative contexts?"* The main scientific contributions of this work include:

- Designing a method that extends users' visible and interactive space through mobile mixed reality telepresence,
- Investigating the effects of the proposed approach in social and collaborative contexts, and
- Proposing a system for "outsourcing" of users' cognitive workload to other users in collaborative applications.

2 BACKGROUND

According to the Cisco report published in 2017 [2], mobile users are the fastest growing group of people who experience and interact with video streaming applications. Therefore, in order to have a higher social relevance and impact, this research focuses mainly on *mobile* mixed reality telepresence systems, taking into consideration the available spatial information (internal and external tracking on mobile devices), and immediacy of mobile interactions.

Similar works can be divided into two categories: systems that focus on telepresence and co-presence experiences, and systems that focus on remote collaboration. In case with telepresence, this work bears resemblance to such projects as JackIn[8], PanoVC[11], and SharedSphere[9], and in case with collaboration this work is similar to "Social Panoramas,"[1] "World-Stabilized annotations,"[4] and Chili[7]. The limitations of presented works include reduced visible and interaction space due to locked-in viewpoints (viewers cannot freely explore a remote location, only seeing what is in streamer's video feed) [4, 7], lack of real-time interactions[1], limited mobility due to semi-tethered (part-mobile, part-stationary) interface[8, 9], and inability to support more than two users simultaneously[1, 4, 7, 8, 11].

3 RESEARCH METHODOLOGY

Based on the observed limitations, we designed a media stream composition method that would allow users to: freely explore a remote location regardless of streamer's viewpoint, share each other's real-time viewpoints, and interact through audio, haptic feedback, and

WOODSTOCK'97, July 1997, El Paso, Texas USA

ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

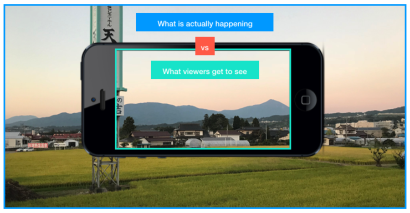

Figure 1: Example of an issue with single-viewpoint video streams: limited visual and interactive space

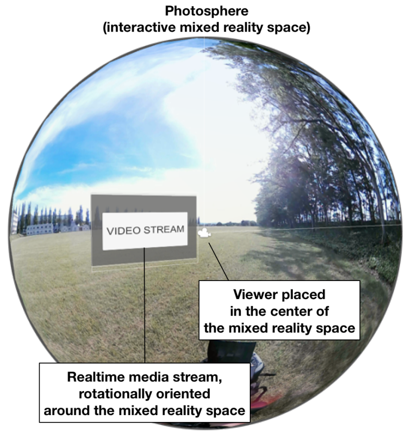

Figure 2: Example of a SLC-based mixed reality space

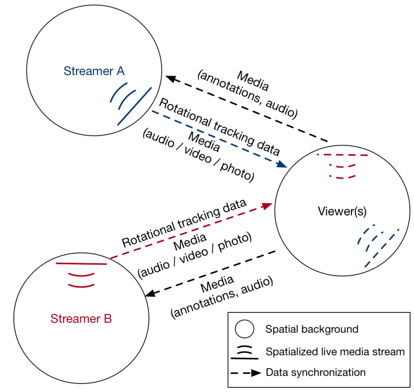

Figure 3: Example of network connection and interaction within the SLC-based applications

realtime annotations. Named as Spatial Livestream Composition (SLC) method, the proposed approach takes advantage of mobile devices' spatial data to create a collaborative mixed reality space. SLC method organizes both spatial and media information into two following categories (Figure 2). *Spatial background* that represents a generalized information source that provides a spatial context in which users can be placed (e.g. a photospherical image or a real-time spherical video stream from a remote location), and *spatialized live media streams* that represent a live media feed coming from a streaming user, spatially oriented within the mixed reality space.

In order to investigate the feasibility of SLC method, we developed a Unity-based framework that allows creating a mixed reality space from web-based photospherical imagery, with rotationally-oriented real-time video streams that represent connected users' viewpoints. All users can interact through live audio and video sharing, as well as through 3D annotations in mixed reality space (Figure 7). The internal rotational tracking is handled through Google Cardboard SDK running in "handheld" (non-VR) mode, whereas network connectivity is implemented through Web Real-Time Communication (WebRTC) protocol, which allows unbounded number of connected users within a single collaborative session (Figure 3).

4 CURRENT RESULTS

The developed framework has been used in three proof-of-concept applications. In the first application [13] we have investigated whether SLC method could be beneficial in collaborative scenarios. We divided users into two groups: streamers (who were physically present in a remote location), and viewers (who were connected remotely). In the experiment we asked viewers to help streamers find an object within a remote location using our SLC-based and non-SLC (single-viewpoint video stream) applications. After each trial we asked users to fill out NASA-RTLX questionnaires, followed by questions on spatial and situational awareness. As a result, we have found that the SLC-based application has shown a statistically significant ($p < 0.05$ via Wilcoxon's Signed Rank Test) decrease in cognitive workload, as well as increase in spatial and situational awareness among viewers in comparison with a "regular" non-SLC application.

In the second application [12, 14] we investigated whether SLC method would make video stream viewing experience engaging in social entertainment context (Figure 4,5,6). We have created an interface similar to Periscope[1], a popular live streaming application, where only streamers can share their realtime video feed, whereas users could interact through audio and by placing "like" buttons within a mixed reality space. In the experiment we asked users to try watching and interacting with live video streams in our application and in Periscope, and answer a questionnaire based on [5]. We have found a statistically significant improvement ($p < 0.05$ via Wilcoxon's Signed Rank Test) in questions on whether users enjoyed interacting with multiple video streams over an average score of 3, but we could not detect statistically significant improvements in questions on whether users preferred our application to Periscope.

5 ONGOING AND FUTURE WORK

In the third application we are currently exploring whether SLC-based applications could be used for "outsourcing" users' cognitive workload in visual search or navigation tasks. The application aims to take advantage of the "instant community" phenomenon that occurs during live video streaming sessions [5]. In such case, the developed application asks connected users to help a streamer either navigate in a remote location, or find a certain object within a spherical photo or video stream. We are currently designing a blockchain-based reward system that would encourage viewers to help streamers in collaborative sessions similar to [10] based on the mixed reality content distribution system that we developed in [15]. In the near future we are planning to expand this application towards remote assistance in accessibility scenarios (for instance, helping visually-impaired users to navigate around locations, or remotely operating a wheelchair similar to [6]).

[1] https://www.pscp.tv/

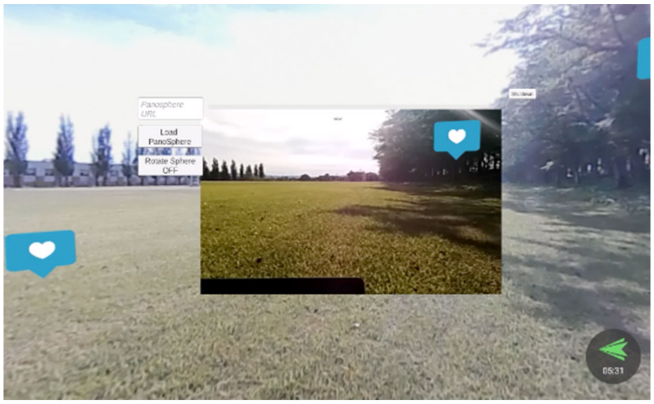

Figure 4: Streamer's view in a SLC-based application

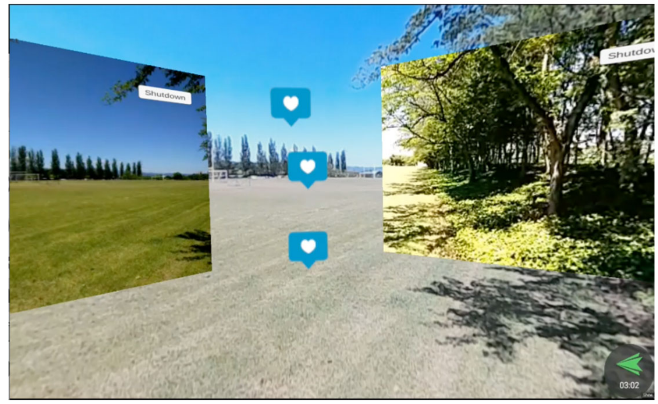

Figure 5: Viewer's view

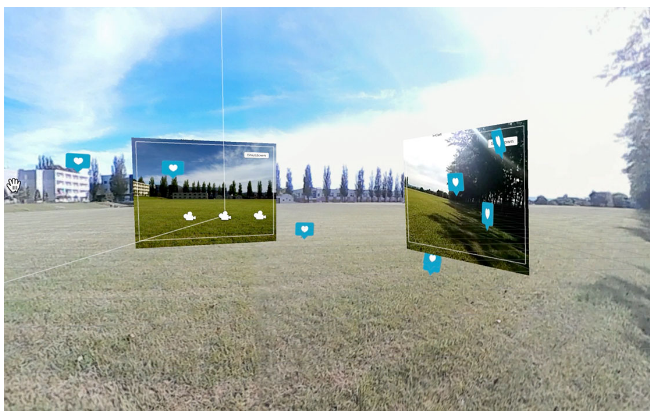

Figure 6: Scene overview

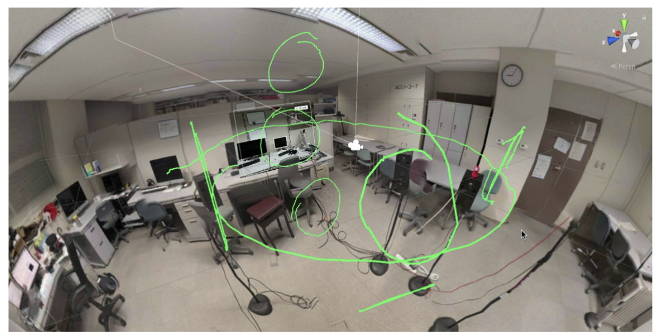

Figure 7: Example of annotations within a collaborative session

REFERENCES

[1] Mark Billinghurst, Alaeddin Nassani, and Carolin Reichherzer. 2014. Social Panoramas: Using Wearable Computers to Share Experiences. In *SIGGRAPH Asia (SA '14)*. ACM, New York, Article 25, 1 pages.

[2] Cisco. 2017. The Zettabyte Era: Trends and Analysis. (Jun. 2017).

[3] Andy Clark and David Chalmers. 1998. The extended mind. *analysis* 58, 1 (1998), 7–19.

[4] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. World-stabilized annotations and virtual scene navigation for remote collaboration. In *Proc. of the 27th annual ACM Symp. on User interface software and technology*. ACM, 449–459.

[5] William A Hamilton, John Tang, Gina Venolia, Kori Inkpen, Jakob Zillner, and Derek Huang. 2016. Rivulet: Exploring Participation in Live Events through Multi-Stream Experiences. In *Proc. of the ACM Int. Conf. on Interactive Experiences for TV and Online Video*. ACM.

[6] Satoshi Hashizume, Ippei Suzuki, Kazuki Takazawa, Ryuichiro Sasaki, and Yoichi Ochiai. 2018. Telewheelchair: The Remote Controllable Electric Wheelchair System Combined Human and Machine Intelligence. In *Proceedings of the 9th Augmented Human Int. Conf.* ACM, New York.

[7] Hyungeun Jo and Sungjae Hwang. 2013. Chili: viewpoint control and on-video drawing for mobile video calls. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1425–1430.

[8] Shunichi Kasahara and Jun Rekimoto. [n. d.]. JackIn Head: An Immersive Human-human Telepresence System. In *SIGGRAPH Asia 2015 Emerging Technologies*. New York.

[9] Gun A. Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. [n. d.]. Sharedsphere: MR Collaboration Through Shared Live Panorama. In *SIGGRAPH Asia 2017*.

[10] Ming Li, Jian Weng, Anjia Yang, Wei Lu, Yue Zhang, Lin Hou, Jia-Nan Liu, Yang Xiang, and Robert H Deng. 2017. CrowdBC: A Blockchain-based Decentralized Framework for Crowdsourcing. *IACR Cryptol. ePrint Arch.* (2017).

[11] Jörg Müller, Tobias Langlotz, and Holger Regenbrecht. 2016. PanoVC: Pervasive telepresence using mobile phones. In *Pervasive Computing and Communications (PerCom)*. IEEE.

[12] Bektur Ryskeldiev, Michael Cohen, and Jens Herder. 2017. Applying Rotational Tracking and Photospherical Imagery to Immersive Mobile Telepresence and Live Video Streaming Groupware. In *SIGGRAPH Asia*.

[13] Bektur Ryskeldiev, Michael Cohen, and Jens Herder. 2018. StreamSpace: Pervasive Mixed Reality Telepresence for Remote Collaboration on Mobile Devices. *IPSJ Journal* (2018).

[14] Bektur Ryskeldiev, Michael Cohen, Jens Herder, and Yoichi Ochiai. 2018 (Accepted). ReactSpace: Spatial-Aware User Interactions for Collocated Social Live Streaming Experiences. In *IEEE Int. Conf. on Systems, Man, and Cybernetics (SMC)*.

[15] Bektur Ryskeldiev, Yoichi Ochiai, Michael Cohen, and Jens Herder. 2018. Distributed metaverse: creating decentralized blockchain-based models for peer-to-peer sharing of virtual spaces for mixed reality applications. In *The 9th Augmented Human Int. Conf.* New York.