QoE Modeling of Delay in 3D Tele-Immersion

| 1st Author Name  Affiliation  City, Country  e-mail address | 2nd Author Name  Affiliation  City, Country  e-mail address | 3rd Author Name  Affiliation  City, Country  e-mail address |
| --- | --- | --- |

# ABSTRACT

Paste the appropriate copyright/license statement here. ACM now supports three different publication options:

* ACM copyright: ACM holds the copyright on the work. This is the historical approach.
* License: The author(s) retain copyright, but ACM receives an exclusive publication license.
* Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single-spaced in Times New Roman 8-point font. Please do not change or modify the size of this text box.

Each submission will be assigned a DOI string to be included here.

## Author Keywords

Telepresence, Delay, Network Performance, QOE.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; See<http://acm.org/about/class/1998> for the full list of ACM classifiers. This section is required.

# INTRODUCTION

Communications technology plays an important role in human development. The invention of telephone made most remote communications instantaneous. From than on, more and more physical meetings were replaced by phone calls, which saves a great deal of time and money. Nowadays, video-mediated telecommunication is becoming popular. It provides convenience for teleconference [2, 3, 4], tele-collaboration [5, 6], presence remotely [7, 8, 9, 10], and so on.

Beyond that, researchers are also exploring telepresence with higher level of immersion. In the last decades, tele-immersion developed rapidly. Several 3D-reconstruction-based telepresence systems were born [11, 12, 13, 14, 15]. They aim at making up for the lack of eye contact, body language and physical presence in video-mediated telecommunications. Microsoft Research’ s Holoportation [15] was quite impressive. They presented an end-to-end system for AR/VR telepresence with high-quality, real-time reconstructions of an entire space. Because of their promising quality of service (QoS) and the fact that hardware devices are getting cheaper and more powerful, we believe that these systems will become practical in the near future.

However, previous works about tele-immersion bias to technological implementations. Only a few study works were conducted on 3D telepresence. Moreover, they either study on specific scenarios [18, 19, 20] or with pseudo-3D systems [2, 16, 17].

We argue that, fundamental studies on mapping quality of service (QoS) to quality of experience (QoE) in 3D tele-immersion is important. We have seen that the industrial standard of telephone contributes to its popularization, i.e., by avoiding network over-engineering [21]. In recent researches, the user experience (UX) studies of video-mediated telecommunications [2, 3, 4, 9, 17] are also helping its improvement. Similarly, an understanding of UX in 3D telepresence may well be helpful to both academic and industrial community.

In this paper, we focus on modeling the impact of delay, which is an important factor of QoS [5], in 3D tele-immersion. We first conducted a large online questionnaire (N=100) to help summarizing the supported tasks of our system. Then, we classified the practical tasks by xx, xx and xx, and selected several typical applications for the user study.

In implementation, we decided not to follow the reconstruction technique with the highest quality [22, 1] proposed by Microsoft Research, but to achieve a more responsive system. The kernel of our system is similar to Maimone et al. ’s work [14]. Thank for the recently development of depth camera (RealSense-D435 was used), GPU (Gtx1080 Ti) and VR device (HTC Vive), the frame rate of our system reach 40fps. Only one frame delay is necessary for the transmission, so the delay can be within 50ms. Furthermore, as several related work mentioned the importance of “shared objects” in 3D telepresence [19], our system was designed to go around the common objects in both sides. This feature indeed extends the supported tasked of our system, for example, one can now play a “piano duet” with a remote friend if they own similar desks (piano in virtual scene).

Result shows that, some tasks with strong interaction require lower latency of 75ms, which breaks the “rule” that 150ms is acceptable for most telecommunication applications [5, 23]. This finding indicates the higher immersion in 3D tele-immersion system compared to video-mediated techniques. On the other hand, some tasks without instantaneous interaction, i.e. playing chess, do not require such a low delay. Because of this variation, we tried and finally found out a prediction from QoS to QoE (R > 80%) based on task classification and QoS/QoE correlations. In fact, I fabricated this paragraph because I do not know the result yet.

# related work

## Tele-Immersion systems

In order to implement a typical tele-immersion system for our studies, we first conducted a brief review. Basically, a 3DTI system requires three processes: reconstruction, transmission and rendering [24]. Finally, we decided to develop our reconstruction algorithm based on TSDF Volume [25] and Marching Cubes [26]. We use network line between computers for high-bandwidth transmission, but do not focus on the transmission part as [27, 19] did. In the studies, we simulated various network performance through software methods. For 3D rendering, we use head-mounted display (HTC Vive).

Below are brief reviews of reconstruction and rendering techniques for 3DTI systems:

### 3D Reconstruction

### 3D Rendering

## Studies in Telepresence

## QoE Measurement

# REFERENCES

1. Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., ... & Tankovich, V. (2016, October). Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (pp. 741-754). ACM.
2. Higuchi, K., Chen, Y., Chou, P. A., Zhang, Z., & Liu, Z. (2015, April). ImmerseBoard: Immersive telepresence experience using a digital whiteboard. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*(pp. 2383-2392). ACM.
3. Nemiroff, G. (1989). Beyond" talking heads": Towards an empowering pedagogy of women's studies. *Atlantis: Critical Studies in Gender, Culture & Social Justice*, *15*(1).
4. Marlow, J., Van Everdingen, E., & Avrahami, D. (2016, February). Taking Notes or Playing Games?: Understanding Multitasking in Video Communication. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 1726-1737). ACM.
5. Donovan, A., Alem, L., Huang, W., Liu, R., & Hedley, M. (2014, September). Understanding How Network Performance Affects User Experience of Remote Guidance. In *CYTED-RITOS International Workshop on Groupware* (pp. 1-12). Springer, Cham.
6. Avellino, I., Fleury, C., & Beaudouin-Lafon, M. (2015, April). Accuracy of deictic gestures to support telepresence on wall-sized displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2393-2396). ACM.
7. Nakanishi, H., Tanaka, K., & Wada, Y. (2014, April). Remote handshaking: touch enhances video-mediated social telepresence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2143-2152). ACM.
8. Misawa, K., & Rekimoto, J. (2015, April). ChameleonMask: Embodied physical and social telepresence using human surrogates. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 401-411). ACM.
9. Rae, I., & Neustaedter, C. (2017, May). Robotic Telepresence at Scale. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 313-324). ACM.
10. Neustaedter, C., Venolia, G., Procyk, J., & Hawkins, D. (2016, February). To Beam or not to Beam: A study of remote telepresence attendance at an academic conference. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (pp. 418-431). ACM.
11. Kurillo, G., Bajcsy, R., Nahrsted, K., & Kreylos, O. (2008, March). Immersive 3d environment for remote collaboration and training of physical activities. In *Virtual Reality Conference, 2008. VR'08. IEEE* (pp. 269-270). IEEE.
12. Petit, B., Lesage, J. D., Menier, C., Allard, J., Franco, J. S., Raffin, B., ... & Faure, F. (2010). Multicamera real-time 3d modeling for telepresence and remote collaboration. *International journal of digital multimedia broadcasting*, *2010*.
13. Maimone, A., & Fuchs, H. (2011, October). Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras. In *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*(pp. 137-146). IEEE.
14. Maimone, A., & Fuchs, H. (2012, October). Real-time volumetric 3D capture of room-sized scenes for telepresence. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2012* (pp. 1-4). IEEE.
15. Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., ... & Tankovich, V. (2016, October). Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (pp. 741-754). ACM.
16. Kuster, C., Ranieri, N., Zimmer, H., Bazin, J. C., Sun, C., Popa, T., & Gross, M. (2012, October). Towards next generation 3D teleconferencing systems. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2012* (pp. 1-4). IEEE.
17. Boustila, S., Capobianco, A., & Bechmann, D. (2015, November). Evaluation of factors affecting distance perception in architectural project review in immersive virtual environments. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology* (pp. 207-216). ACM.
18. Benko, H., Jota, R., & Wilson, A. (2012, May). MirageTable: freehand interaction on a projected augmented reality tabletop. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 199-208). ACM.
19. Beck, S., Kunert, A., Kulik, A., & Froehlich, B. (2013). Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics*, *19*(4), 616-625.
20. Pejsa, T., Kantor, J., Benko, H., Ofek, E., & Wilson, A. (2016, February). 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* (pp. 1716-1725). ACM.
21. Bergstra, J. A., & Middelburg, C. A. (2003). ITU-T Recommendation G. 107: The E-Model, a computational model for use in transmission planning.
22. Dou, M., Khamis, S., Degtyarev, Y., Davidson, P., Fanello, S. R., Kowdle, A., ... & Kohli, P. (2016). Fusion4d: Real-time performance capture of challenging scenes. *ACM Transactions on Graphics (TOG)*, *35*(4), 114.
23. ITU-T, I. T. U. T. (2003). Recommendation G. 114. *One-Way Transmission Time, Standard G*, *114*.
24. Fuchs, H., State, A., & Bazin, J. C. (2014). Immersive 3d telepresence. *Computer*, *47*(7), 46-52.
25. Curless, B., & Levoy, M. (1996, August). A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (pp. 303-312). ACM.
26. Lorensen, W. E., & Cline, H. E. (1987, August). Marching cubes: A high resolution 3D surface construction algorithm. In *ACM siggraph computer graphics* (Vol. 21, No. 4, pp. 163-169). ACM.
27. Pece, F., Kautz, J., & Weyrich, T. (2011, September). Adapting standard video codecs for depth streaming. In *Proceedings of the 17th Eurographics conference on Virtual Environments & Third Joint Virtual Reality* (pp. 59-66). Eurographics Association.