<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *18th International Conference on Human-Computer Interaction (HCI International), Toronto, July 17-22, 2016.*

Citation for the original published paper:

Khan, M S., Li, H., ur Réhman, S. (2016)

Tele-Immersion: Virtual Reality based Collaboration

In: Constantine Stephanidis (ed.), *HCI International 2016: Posters' Extended Abstracts : 18th International Conference, HCI International 2016, Toronto, Canada, July 17-22, 2016, Proceedings, Part I* (pp. 352-357). Springer

Communications in Computer and Information Science

https://doi.org/10.1007/978-3-319-40548-3_59

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-120461>

Tele-Immersion: Virtual Reality based Collaboration

M.S.L. Khan¹, Haibo Li², and Shafiq Ur Réhman¹

¹ Department of Applied Physics and Electronics, Umeå University, 901 87, Sweden.

² KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden.

`muhammad.sikandar.lal.khan@umu.se`

Abstract. The ‘perception of being present in another space’ during video teleconferencing is a challenging task. This work makes an effort to improve upon a user perception of being ‘present’ in another space by employing a virtual reality (VR) headset and an embodied telepresence system (ETS). In our application scenario, a remote participant uses a VR headset to collaborate with local collaborators. At a local site, an ETS is used as a physical representation of the remote participant among his/her local collaborators. The head movements of the remote person is mapped and presented by the ETS along with audio-video communication. Key considerations of complete design are discussed, where solutions to challenges related to head tracking, audio-video communication and data communication are presented. The proposed approach is validated by the user study where quantitative analysis is done on *immersion* and *presence* parameters.

Keywords: Tele-immersion, Virtual reality, Embodied telepresence system, presence, distal attribution, spatial cognition

1 Introduction and Background

Virtual Reality (VR) headsets are becoming more popular due to cheaper prices, more immersive experience, high quality, better screen resolution, low latency and better control. According to statistics, around 6.7 million people have used VR headsets in year 2015 and it is expected to grow to 43 million in year 2016 and 171 million in year 2018 [1]. The most valuable application for these headsets could be video teleconferencing to enhance a personal and social experience by more ‘immersive’ and ‘presence’ feelings [3]. Among other problems, for example, gaze perception and nonverbal communication; ‘different-environmental-spaces’ poses a big problem in standard video teleconferencing [4]. This problem leads to a lack of immersiveness and presence experience in standard conferencing setup. Immersion is a key feature of virtual reality headsets. Immersion is a phenomenon of *getting lost, involved or drawn into something*. Teleimmersion is when *geographically distant located participant immerse himself to be located in a new environment, where he shares a same space with his collaborators* [3].

One-to-many video teleconferencing setups are quite common in many applications, where *one* person (called a ‘remote user’) collaborates with his/her

many collaborators (called a ‘local collaborators’) using audio-video modalities. The standard video teleconferencing setups pose certain problems at remote user end. The remote user cannot i) take turns, ii) do side-conversation, iii) make an eye-contact, iv) explore the distant environment, etc. [4, 5]. This affects a conferencing experience of the remote person which results in lack of participation, immersion, engagement and involvement in in-meeting discussions [7].

Over the years, number of solutions have been proposed which varies from software algorithms to hardware platforms. The software solutions apply computer graphics, image processing, computer vision and augmented reality techniques to solve the above mentioned problems in standard video teleconferencing [8]. The most common example of hardware based solution is mobile robotics telepresence (MRP) systems. The basic construction of almost all MRPs consist of a mobile robotic base, an LCD screen, a camera, a microphone and some non-verbal gestures, like hand gestures, etc. The review on MRPs can be found in [7]. The other types of hardware solutions can be broadly categorized into i) an anthropomorphic design and ii) a non-anthropomorphic design. The characteristics and appearance of anthropomorphic video teleconferencing systems (AVTS) are similar to real-human characteristics and appearance. On the other hand, the appearance of non-anthropomorphic video teleconferencing systems (NAVTS) are not like human but they still possess some human skills. The review on AVTS and NAVTS can be found in [4].

Despite all these solutions, a remote person still feels *disconnected* with his local collaborators. The major reason is that the above mentioned solutions did not address an important issue of ‘*different environmental spaces*’ between the remote and local participants. This work aims to provide a solution to minimize the distance between the remote person and the local collaborators.

2 System Overview

The system diagram of our virtual reality (VR) based video teleconferencing setup is shown in Figure 1. As compared to a traditional conferencing setup, our setup employs a virtual reality headset at a remote-user end. This work uses a google-cardboard [2] in combination with smartphone, but it can be extended to other VR devices. The VR headset displays local environment/participants to the remote user.

The right side of Figure 1 shows our embodied telepresence system (ETS), which is placed among the local collaborators. The ETS is 3-DOF neck robot which actuates an attached tablet-PC. The tablet-PC is used for audio-video communication and a neck robot is used for presenting head gestures. The ETS embodies the remote user’s head movements at the local site. The details of ETS can be found in [6].

We have proposed two head tracking algorithms to map the remote person’s head movements to ETS mechatronic platform. Our first solution uses external IMU sensor to record 3 orientation angles, i.e., yaw, pitch and roll angles [9]. Our second solution uses an orientation sensor of a smartphone to capture these

Fig. 1: Virtual Reality based video teleconferencing setup: On the left, we have a remote site and on the right we have a local site. The remote person uses VR headset (in our case google cardboard with mobile phone). Remote person is represented by our Embodied Telepresence System (ETS) at the local site. The ETS embodies the remote person head with a mechatronic robot where tablet-PC is used for audio-video communication. The audio-video communication is implemented through WebRTC. The head tracking of remote person is done by i) orientation sensor of mobile phone and/or ii) external IMU sensor. The data communication is done through i) Xbee for short range and/or ii) client-server application for long range.

head rotations. The data communication between the remote site and local site is done by using i) a XBEE for short range communication and ii) a client-server application for long range communication.

The final step in our design is an audio-video communication. The traditional conferencing software (e.g. skype) cannot be used here as these software does not allow a stereo view for VR headsets. Therefore, we have proposed our own audio-video communication technique, which is based on webrtc [13]. Our webrtc based solution is a peer-to-peer connection between the remote and local sites, and the local video can be edited to make it a suitable input for a VR headset.

3 User Study: Remote User Perspective

A user study is conducted to compare our proposed virtual reality (VR) based video teleconferencing setup with the setup given in [9]. The setup in [9] uses embodied telepresence system (ETS) and the study with this setup shows promising results as compared to standard video teleconferencing softwares (e.g. skype).

3.1 Participant and Procedure

Seven participants, aging from 18-40 were recruited from the campus of Umeå University-Sweden, including both the students and staff members. For each

round of an experiment we have one remote person and six local participants (one-to-many video teleconferencing scenario). The remote participant is asked to collaborate with local participants/environment based on a flexible-script provided by us. There are two conferencing scenarios for each remote person; collaboration performed with i) the ETS-only setup (ETS) [9] and ii) our proposed VR in-combination with ETS setup (VR-ETS). At the end of the experiment, the participants are asked questions related to an immersion and presence parameters, attention, mental construction, emotional state, engagement, ease-of-interaction and being-there [10, 11].

- **Attention:** Which system increases the attention level during collaboration?
- **Mental Construction:** Which system helps you in creating a mental representation of a local environment?
- **Emotional State:** Which system has more affect on an emotional state of a person?
- **Engagement:** Which system makes you believe that your actions are inter-dependent, connected to and responsive to the others?
- **Ease of Interaction:** Which system provides a better quality of experience (QoE)?
- **Being-there:** Which system creates a sense of spatial presence in a local environment?

We have used a likert style 7-point rating system (1 to 7) [12] to compared ETS with VR-ETS, where 1 represents strong disagreement (negative) and 7 represents strong agreement (positive).

3.2 Results and Discussion

The comparative results (ETS vs VR-ETS) of above mentioned questions are shown in Figure 2. These results are presented in the form of means and standard-deviations of participant responses. The blue bars show the responses for the ETS and the red bars show the responses for the VR-ETS. The results clearly show that VR-ETS outperforms ETS on all immersion and presence parameters. Among all parameters, *being-there* got the highest score for the VR-ETS. This is because of immersion capability of a VR headset. Furthermore, the quality of interaction and level of engagement is also improved with an introduction of the VR headset. The movement capability of ETS in combination with a VR headset helps a remote user to get aware of the local environment and increases his level of attention which in-turn affects the emotional state of the remote person.

4 Conclusion

We have used a virtual reality (VR) headset in combination with our previously developed Embodied Telepresence system (ETS) for a novel video teleconferencing experience. The use of the VR headset disconnects the remote person

Fig. 2: Comparative results of the ETS vs the VR-ETS on immersion and presence parameters. The participants' response are presented in the form of mean and stad. dev. on a 7-point likert scale. The blue bars show responses for the ETS and the red bars show responses for the VR-ETS.

from his space and makes him/her feel-like to be at the local collaborator space. Furthermore, the ETS facilitates the remote person to take turns, do side conversation, explore the local environment and specify his gaze, etc. The user study shows that our setup increases the *immersiveness* and *feeling-of-presence* experience of a remote person at a local site. The limitation with current system is that the VR headset occludes significant portion of a remote person's face. This issue will be addressed in future work.

References

1. Virtual reality users, website, <http://www.statista.com/statistics/426469/active-virtual-reality-users-worldwide/> (2016)
2. Google Cardboard, Website, <https://www.google.com/get/cardboard/> (2016)
3. Perry, Tekla S.: 'Virtual reality goes social.' *Spectrum*, IEEE 53.1 (2016): 56-57.
4. Khan, M. S. L., and S. Rehman.: *Distance Communication: Trends and Challenges and How to Resolve them. Handbook: Strategies for a Creative Future with Computer Science, Quality Design and Communicability.* Blue Herons (EDs) (2014).
5. K. Fischer and T. Tenbrink.: *Video conferencing in a transregional research cooperation: Turn-taking in a new medium.* Technical report, <http://nats-www.informatik.uni-hamburg.de/~fischer/VKfischertenbrink.pdf>.
6. Khan, Muhammad Sikandar Lal, Haibo Li, and Shafiq Ur Rhman.: *Embodied Tele-Presence System (ETS): Designing Tele-Presence for Video Teleconferencing. Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments.* Springer International Publishing, 2014. 574-585.
7. Kristoffersson, S. Coradeschi, A. Loutfi: *A Review of Mobile Robotic Telepresence Advances in Human- Computer Interaction*, Volume 2013, 17 pages, 2013.

8. P. Eisert: Immersive 3D Video conferencing: challenges, concepts, and implementations, Proc. VCIP 2003, pp. 6979.
9. Khan, Muhammad Sikandar Lal, et al.: A pilot user's prospective in mobile robotic telepresence system. Asia-Pacific Signal and Information Processing Association, 2014 Annual Summit and Conference (APSIPA). IEEE, 2014.
10. Heeter, Carrie.: Being there: The subjective experience of presence. Presence: Teleoperators Virtual Environments 1.2 (1992): 262-271.
11. Diemer, Julia, et al.: The impact of perception and presence on emotional reactions: a review of research in virtual reality. Frontiers in psychology 6 (2015).
12. Likert, Rensis.: A technique for the measurement of attitudes. Archives of psychology (1932).
13. Johnston, Alan B., and Daniel C. Burnett.: WebRTC: APIs and RTCWEB protocols of the HTML5 real-time web. Digital Codex LLC, 2012.