AVATAREX: Telexistence System based on Virtual Avatars

Timo Koskela[1], Mounib Mazouzi[1], Paula Alavesa[1], Minna Pakanen[1],
Ilya Minyaev[1], Eero Paavola[2], Jere Tuliniemi[2]
Center for Ubiquitous Computing,
FI-90014 University of Oulu
Finland
[1]firstname.lastname@oulu.fi, [2]firstname.lastname@student.oulu.fi

ABSTRACT

The telexistence technology can provide many kinds of benefits for the society. These include new ways of remote work, empowerment of handicapped and elderly people, and creation of new immersive and environmentally-friendly forms of tourism, travel, shopping, sports and leisure time activities. In this paper, we introduce AVATAREX, a telexistence system based on virtual avatars. AVATAREX provides means for connecting users that are simultaneously occupying the same space in the real world and its virtual replica. Using an indoor prototype implementation of AVATAREX and a simple collaborative game, we investigated how users experience co-presence in a telexistence system based on virtual avatars and measured the performance of AVATAREX on high-end smart glasses. Based on our findings, users wearing virtual reality gear reported a stronger sense of co-presence compared to users wearing augmented reality gear. Unexpectedly, users wearing smart glasses reported a lower sense of co-presence than users using a tablet for augmented reality experience.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; *Empirical studies in HCI; Interaction paradigms*; Mixed / augmented reality; Virtual reality; Collaborative interaction

KEYWORDS

virtual reality, augmented reality, mixed reality, user study

ACM Reference format:

T. Koskela, M. Mazouzi, P. Alavesa, M. Pakanen, I. Minyaev, E. Paavola, and J. Tuliniemi. 2018. AVATAREX: Telexistence System based on Virtual Avatars. In *Proceedings of ACM Augmented Human conference, Seoul, Korea, February 2018 (AH'2018)*, 8 pages.
DOI: xx.xxxx/xxxxxxx.xxxxxxx

AH2018, February 7–9, 2018, Seoul, Republic of Korea

ACM ISBN xxx-x-xxxx-xxxx-x/xx/xx
https://doi.org/xx.xxxx/xxxxxxx.xxxxxxx

1 INTRODUCTION

The concept of telexistence was first introduced in 1980 by Tachi [33]. According to Tachi [31], telexistence "*refers to the general technology that allows a human being to experience a real-time sensation of being in a place other than his/her actual location and to interact with the remote environment, which maybe real, virtual or a combination of both*". Telexistence is often connected with telepresence [21], which is a very similar concept, but with a narrower scope. The most significant difference between the two concepts is that telepresence does not include telexistence in virtual environments or telexistence in a real environment through a virtual environment.

The telexistence technology can provide many kinds of benefits for the society. These include, but are not limited to, new ways of remote work, empowerment of handicapped and elderly people, and creation of new immersive and environmentally-friendly forms of tourism, travel, shopping, sports, and leisure time activities [14, 25, 32]. The traditional telexistence systems, such as *TELESAR* [32], have been implemented using remotely operated robots housing an array of different types of displays and sensors. With the introduction of contemporary virtual reality (VR) and augmented reality (AR) gear, it is currently possible to implement telexistence systems using virtual avatars instead of robots [2, 13, 24]. These two technological approaches are not mutually exclusive, but they rather complement each other. Robots are capable of physical interaction with the environment and are capable of operating in environments that would be hazardous for people not wearing specialized gear [32]. However, robots are very expensive and their visual appearance is not easily modifiable. In addition, their movement may be restricted to certain kinds of terrain. Virtual avatars, in turn, are cheap and their visual appearance can be easily modified. They are not restricted to a certain kind of terrain, but can interact with the environment only through digital means.

Over the last years, massive R&D investments [30] have led to the introduction of novel VR/AR gear. As a result, also research on telexistence using virtual avatars has started to gain momentum in the academic community. During the last few years, different kinds of collaborative systems have been proposed that utilize virtual avatars for enabling telexistence (e.g. [2, 7, 13, 19, 24, 25]). However, these systems are typically limited to specific indoor spaces and are capable of supporting only a small number of users. In addition, a few telexistence systems do already exist [24, 25] that connect users between indoor and outdoor spaces. In these systems, indoor users are present in the virtual replica of the shared space, whereas outdoor users are present in the physical location of the shared space.

Until today, outdoor users have communicated with indoor users through a tablet or a smartphone which enable limited means for interaction and a sense of presence compared to the contemporary smart glasses (or AR glasses) [23]. However, the use of smart glasses also brings forth many technical, social and safety-related challenges, which need to be addressed. Some of these are the same as with today's mobile devices, while some are still undiscovered. Technical challenges include unreliability of positioning [14], limited bandwidth and high latencies of wireless networks [10, 11], limited battery life as well as motion capture outside laboratory environments, whereas social and safety challenges are related to vacancy problem [17], self-expression, visibility, interaction, collision, and control [6, 12, 14].

In this paper, we introduce AVATAREX, which is a telexistence system based on virtual avatars. AVATAREX provides means for connecting users that are simultaneously occupying the same space in the real world and its virtual replica. The AVATAREX system follows the principles of hybrid reality [28] or dual reality [17] in which the virtual world is enriched with data sensed from the real world, but also data originating from the virtual world is reflected to the real world using AR technologies. We have implemented a fully functional indoor prototype of the AVATAREX system that supports use of both mobile devices and smart glasses for AR experience. For investigating how users experience co-presence in a multiplatform telexistence system based on virtual avatars, we implemented a simple collaborative game for AVATAREX. This game was played by 16 participants who provided new insights on how the sense of co-presence could be developed in telexistence systems based on virtual avatars. Additionally, for identifying the performance bottlenecks with the current smart glasses, we examined the technical performance of the AVATAREX prototype.

The rest of the paper is organized as follows. Section 2 presents the related work. Section 3 describes the AVATAREX system in detail, and Section 4 presents the prototype implementation of AVATAREX. Section 5 presents the user study and the technical evaluation setups, whereas Section 6 presents the results. Finally, Section 7 concludes the paper.

2 RELATED WORK

In this section, some of the recent collaborative systems enabling telexistence are presented. In addition, previous work on presence, co-presence, and immersion is examined.

2.1 Telexistence

Several telexistence systems have been presented in prior research. Most of them rely solely on VR. In [25], the 3DLive platform is developed that enables telexistence in the area of sport applications. With 3DLive, people doing the same type of sports both indoors (e.g. running on a treadmill) and outdoors (e.g. running in the city center) can share the experience. Users can communicate via voice. In addition, the location of the outdoor user is tracked using GPS and visualized in a virtual environment for the indoor user. However, no visualization is provided for the outdoor user. In [2], an indoor telepresence system is presented that allows distributed groups of users to meet in a shared virtual environment. The system uses two projection-based multi-user 3D displays to enable collaboration. In addition, a cluster of depth cameras are required at each site for continuous capturing of participants and any other physical objects. In [15], an indoor telepresence platform is introduced which can be used for remote lectures, remote operations, or immersive live broadcasting of events. With the platform, a user at a remote location can share and exchange emotional expressions with other users at the location of the real world event by using 360° cameras, environmental sensors compliant with MPEG-V, and a game cloud server combined with holographic display technology. In [24], a Beaming Scene Service (BSS) is introduced which allows integrating heterogeneous systems through process coordination. The paper also presents a telexistence application in which BSS has been utilized. In this application, indoor and outdoor users can explore the same campus area and participate in a scavenger hunt. The application connects indoor and outdoor users both in a virtual environment based on Second Life and in a 2D map interface.

Prior research has also shown benefits of AR in telexistence. In [13], an AR-based telexistence application is presented. The application utilizes virtual avatars whose motion is adapted to the physical configuration of the remote site. For capturing the users' motion, Microsoft Kinect is used. In [19], a modular network solution is presented that enables multi-party real-time communication for telexistence applications using virtual avatars. The application supports both use of natural and synthetic avatars. Finally, in [7], a recent telexistence system is introduced. This system enables high visual quality and real-time 3D reconstruction of an entire space including people, furniture and all other objects using custom-made depth cameras. The system supports both VR and AR users.

Many of the previous telexistence systems are either limited to indoor spaces as they require dedicated hardware unsuitable to be used outdoors or they do not take the advantage of novel AR technologies. In this paper, our goal is to particularly examine how users sense the level of co-presence in a telexistence system when both VR and AR technologies are used.

2.2 Presence, Co-presence and Immersion

Presence, the sense of being there and a memory of there as a place, is an elusive target for measurement as it can be considered an everyday experience and a basic property of human conscious experience [4, 26, 27]. Co-presence is the sense of being there with another person and social presence is a sense of intimacy and salience of the connection to another person. For a person to have a sense of co-presence, the participants need to interact or feel that they are interacting with a fellow person [20, 22]. Co-presence can also be defined more concisely as a mere sense of others, one aspect in social presence, which is a more holistic term for social connectedness [5, 9]. In this paper, we consider co-presence as it is described by Goffman [8] and measured by Novak and Biocca [22], *a distinctly measurable sense of being together inside a media, where people have a sense of perceiving and being perceived by others.* Regardless of the definition of co-presence, a system designed for telexistence needs to facilitate communication and have some degree of agency build in for the users.

In the context of telexistence, co-presence can be seen as a more valid and measurable target than presence. Higher level of immersion has been observed to result in a higher sense of presence [4]. However, the connection between co-presence and immersion using modern VR and AR gear has not been thoroughly explored, although some previous studies suggest that there is no positive correlation between immersion and co-presence [29]. Preliminary user study conducted with a recent telexistence system [7] suggests that the system can provide a high sense of co-presence. This user study along with others confirms that the visual quality of avatars affects the sense of co-presence [5][7]. Based on our best knowledge, there is no previous work that specifically focuses on measuring co-presence in a multiplatform telexistence system that connects both VR and AR users.

3 AVATAREX

Our aim was to facilitate telexistence for people using both VR and AR technologies. Therefore, the AVATAREX system is somewhat complex and deserves a detailed description.

3.1 Overall System Architecture

The core components of the AVATAREX system architecture are (1) the AR client and (2) the VR client, which communicate with each other using (3) a network server as shown in Figure 1.

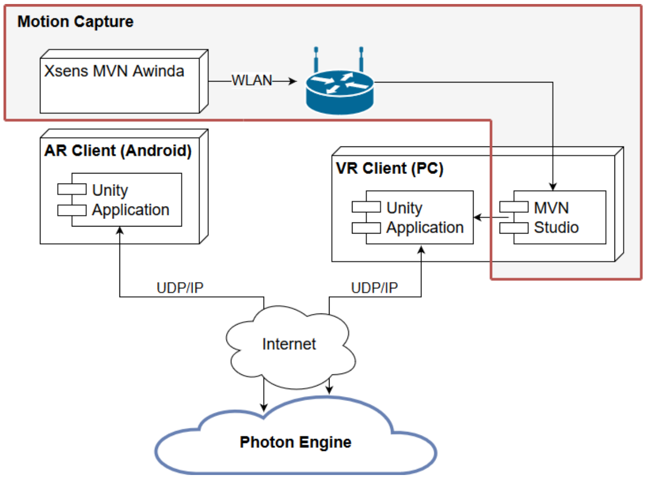


Figure 1: The AVATAREX system architecture.

The AR client runs on Android devices including smartphones, tablets, and smart glasses. With smart glasses, there is also an option to use a motion capture suit Xsens MVN Awinda for creating a fully animated virtual avatar presentation of the AR client in real-time (see "Motion Capture" in Figure 1). For communicating the avatar animation to the VR client, the PC hosting the VR client must also run a proprietary software component called MVN Studio. As a network solution, Photon Engine (www.photonengine.com) was used. In addition to message brokering, it implements voice communication between the AR and VR clients.

3.2 VR Client

The software for the VR client was implemented as a Unity application. In the implementation, Unity plugins for SteamVR, Photon and Xsens were used. The SteamVR plugin extends Unity with components that simplify the setup required for an application to be VR compatible, the Photon plugin enables the network communication, and the Xsens plugin acts a middleware between MVN Studio and Unity. For the Unity application, also a few custom components were implemented including AvatarController that moves and rotates the VR avatar based on user inputs; and NetworkManager that handles network communication by implementing some of the Photon related callback functions. For the VR client, also a 3D model of a room was implemented and used in the user study.

3.3 AR Client

In Figure 2, the software architecture of the AR client is illustrated. In the Unity application, four different plugins were used. Augumenta Unity plugin was used for hand gesture recognition, Kudan Unity plugin was used for detecting predefined markers, EveryPlay plugin was used for recording the view of the AR client, and Photon plugin was used for network communication. Augumenta Unity plugin uses Augumenta Android plugin to get the camera feed from the Android OS. This camera feed is further sent to Kudan Unity plugin for marker detection.

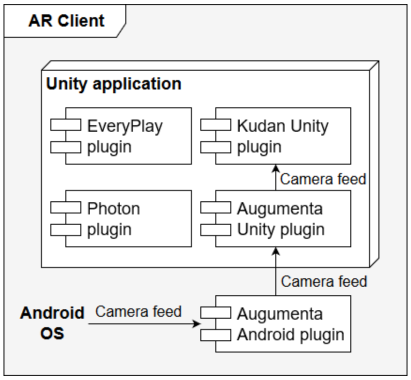


Figure 2: The software architecture of the AR client.

In the prototype implementation of AVATAREX, markers were used both for (1) determining the position and orientation of the AR user; and (2) for placing the VR user's avatar in the view of the AR user. It should be noted that besides markers any other suitable tracking technology could also be used. In outdoors, more feasible approach would be, for instance, to use point clouds based tracking technology [16].

For the Unity application, also a few custom components were implemented. These include ARMovement that calculates the position and orientation of AR user's device based on the position and rotations of the marker; and VRMovement that receives the positions and rotations of the VR user and transfers them into the virtual avatar rendered on top of the marker for the AR user.

3.4 Networking

Networking in the prototype implementation of AVATAREX is enabled by Photon Engine that consists of several different services. For AVATAREX, services used were Photon Unity Networking (PUN), Photon Realtime and Photon Voice. PUN is a plugin for Unity that makes it possible for AR and VR clients to connect to the other Photon services. Photon Realtime is a server solution and Photon Voice is an add-on (requiring PUN) that enables voice communication between AR and VR clients.

Photon services are configured and hosted in the Photon Cloud. Once a Photon service has been configured, an AppID is assigned for that particular service. Unity applications can utilize a Photon service by adding its AppId to the configuration of the PUN plugin. This way, PUN connects both AR and VR clients to Photon Realtime within the Photon Cloud.

Photon Realtime uses a room-based architecture for handling multiplayer sessions. Any kind of communication between clients when not in a room, is not possible. In addition, a client can only be in one room at a time. The room-based architecture matches well with the requirements of AVATAREX as only the users occupying the same place (either in the physical world or in the virtual replica) can communicate with each other.

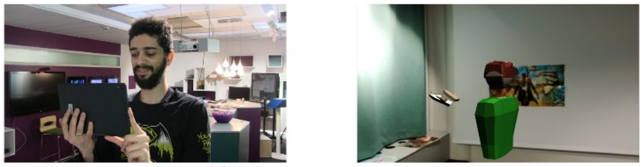
4 PROTOTYPE IMPLEMENTATION OF AVATAREX

4.1 Hardware

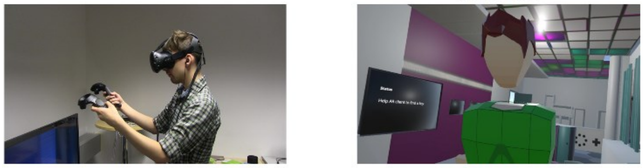
In the prototype implementation of AVATAREX, HTC Vive head-mounted display (HMD) was used as VR equipment. This decision was made based on our experiences that the tracking of the user and her/his handheld controllers was more robust with HTC Vive than any other VR HMD equipment we have tested so far. As AR equipment, ODG R7 smart glasses (2016) and Nexus 9 tablet (2014) were used. ODG R7 is equipped with a capable Snapdragon 805 2.7GHz quad-core processor and 3GBs of RAM. Nexus 9 is equipped with Denver 2.3GHz dual-core processor and 2GBs of RAM. For full-body motion capture, a combination of Xsens MVN Awinda and Link was used. Our setup had 17 motion trackers that were connected to a wireless transmitter using flexible cables. The wireless transmitter was attached to the back of the AR user.

4.2 User Interface

The user interface (UI) in AVATAREX is different for AR and VR users. For the AR user, an abstract humanoid avatar of the VR user was created using Blender (see "AR view" in the upper right corner in Figure 3 and Figure 4). It consists of the upper body with a head and two hands, all parts being separated from each other. The body and head move and roll according to the position and rotation of HTC Vive, while hands duplicate the position and rotation of handheld controllers.


AR user with a tablet

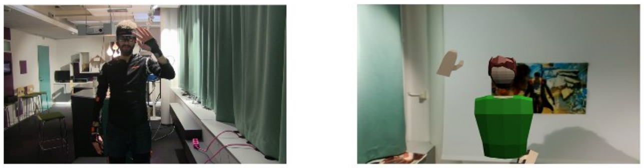
AR view


VR user

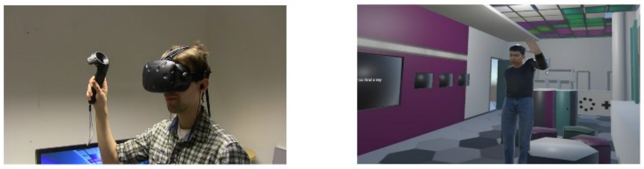
VR view

Figure 3: The avatars implemented for AVATAREX without motion capture.

For the VR user, we implemented two different avatars of the AR user. First, when the AR user was equipped only with a mobile device or smart glasses, the same abstract humanoid avatar was used without any animations (see "VR view" in the lower right corner in Figure 3). Second, when the AR user was equipped with smart glasses and the motion capture suit, fully animated human avatar created with Autodesk character generator (see "VR view" in the lower right corner in Figure 4) was used. In addition to the avatar representing the AR user, the UI for the VR user includes the virtual replica of the shared real-world place.


AR user with smart glasses and mocap

AR view


VR user

VR view

Figure 4: The avatars implemented for AVATAREX with motion capture.

5 THE USER STUDY

The aim of the user study was to investigate how users experience the co-presence in the AVATEREX system. In this paper, we consider co-presence as it is described by Goffman [8] and measured by Novak and Biocca [22], *a distinctly measurable sense of being together inside a media, where people have a sense of perceiving and being perceived by others.*

5.1 Participants

For the user study, we had altogether 16 participants whose age varied between 23 and 41 years with the mean age of 28. Eleven of the participants were male. All participants owned a smartphone, but only six of them owed a tablet device.

Regarding the participants' familiarity with 3D technologies, the following information was discovered. Ten of the participants had played 3D games before. Furthermore, thirteen of the participants had played games that were controlled by hand and/or body movements. Six of the participants were also familiar with 3D virtual worlds such as Second Life or World of Warcraft. Ten of the participants had used VR HMD equipment before, but only two of the participants had prior experience in AR applications or hardware. Finally, twelve of the participants had used devices that allowed them to experience 3D content.

5.2 Test Setup and Use Scenario

For investigating how users experience co-presence in a telexistence system based on virtual avatars, we implemented a simple collaborative game for AVATAREX. We used a large room as the gaming area. An accurate 3D model of this room was created for the VR user as illustrated in Figure 5. It is important to notice that the VR user shared the same gaming area, but was physically in another room than the AR user.

Figure 5: The gaming area (left) and its virtual replica (right).

In our study, the use scenario was a simple task that required constant collaboration between the AR and VR users. The participants could instruct each other through audio connection. They could also point and gesture with their avatars to enhance communication. Both users could also move freely in their small designated areas. The use scenario consisted of two phases:

1. The VR user guides the AR user to find a key that is only visible for the VR user. For guidance, the VR user can use voice, his/her position and hand movements. After reaching the correct position, the AR user must make a predefined gesture with his/her hands to take the key.
2. The AR user guides the VR user to use the key to open a chest that is only visible for the AR user. For guidance, the AR user can use voice, his/her position and hand movements (when the motion capture suit is used). After reaching the correct position, the VR user must touch the chest with the key to open it.

In the game, the location of the key and the chest was varied between participating pairs. For the user study, we had two different clients (AR and VR) and two different versions of the AR client (mobile device only shown in Figure 3 and smart glasses with the motion capture suit shown in Figure 4). As a result, we had altogether four different conditions to investigate:

1. AR mobile (ARM)
2. VR mobile (VRM)
3. AR motion capture (ARMC)
4. VR motion capture (VRMC)

Eight of the 16 participants experienced the ARM and VRM conditions while the other eight experienced the ARMC and VRMC conditions. The reason for not trying out all the conditions with all participants was the relatively long time that was required for wearing and initializing the motion capture suit as the sensors were attached to different body parts using individual straps. Furthermore, we wanted to minimize the effect of recollection bias in the results.

5.3 Data Collection

Before the actual user study, a pilot study was conducted with a few members of the research staff. The aim was to refine the study design and rehearse the flow of the use scenario.

In the beginning of the actual user study, participants were first asked to fill in a background questionnaire. In addition to the basic demographic information, we collected more detailed information about participants' familiarity with different 3D technologies (see Section 5.1. for more details).

The participants were testing the conditions in pairs. First, one of them acted as the AR user (see Figure 6) and the other one as the VR user. Therefore, the experiment was automatically counterbalanced. After completing the task, they filled in the co-presence questionnaire that was a modified version of the questionnaire introduced in [3]. The used questionnaire contained the following items that were evaluated using a 7-point Likert scale:

Q1. To what extent did you have a sense that the other participant was in the same place as you during these events?
Q2. To what extent did you have a sense that you were in the same place as the other participant during the experience?
Q3. To what extent did you have a sense of the presence of the other participant during these events?
Q4. To what extent did you have a feeling that you were collaborating with a real person?
Q5. When you think back about your last experience, do you remember this as more like talking to a computer or communicating with a person?
Q6. To what extent did you have a sense of being "part of the group"?

The options for answers ranged from *never* to *all the time*. After filling in the questionnaire for the first condition, the participants switched the roles, completed the task again and filled in the same co-presence questionnaire for the second condition.

Figure 6. A test user with ODG R7 smart glasses and Xsens motion capture suit (ARMC).

5.4 Performance Measurements

For identifying the potential performance bottlenecks with the current smart glasses, we also examined the technical performance of the AVATAREX prototype. For recording the different performance indicators including (1) CPU load, (2) GPU load and (3) memory consumption, we used Trepn Power Profiler that has been developed for Android devices with a Qualcomm Snapdragon processor. During the 70 seconds of recording, the abstract humanoid avatar was shown on the screen of the Nexus 9 tablet and the ODG R7 smart glasses without any animations. The first 10 seconds of the recordings were omitted as the performance indicators fluctuated due to the application launch. Finally, we also recorded (4) FPS using an Android application GameBench.

6 RESULTS

In this section, the results for both user study and technical performance evaluation are presented.

6.1 User Study: Co-presence

We conducted non-parametric Kruskal Wallis test (also known as one-way ANOVA) [18] for the answers of our co-presence questionnaire. Only the first three items in the co-presence questionnaire gave statistically significant results (see Table 1 and Table 2). The last three questions omitted inquire about the sense of having another person there realistically represented. Therefore, the lack of statistical significance might suggest that the avatar realism needs to be further improved to get more conclusive results. Rather small sample size (N=8 per condition) may also have affected the results concerning the last three questions.

Table 1: Co-presence questionnaire results.

Question	Condition	N	Mean Rank
Q1: To what extent did you have a sense that the other participant was in the same place as you during these events?	ARM	8	14,56
	VRM	8	18,56
	ARMC	8	10,00
	VRMC	8	22,88
	Total	32	
Q2: To what extent did you have a sense that you were in the same place as the other participant during the experience?	ARM	8	12,81
	VRM	8	18,50
	ARMC	8	10,94
	VRMC	8	23,06
	Total	32	
Q3: To what extent did you have a sense of the presence of the other participant during these events?	ARM	8	13,56
	VRM	8	18,50
	ARMC	8	10,94
	VRMC	8	15,06
	Total	32	

Table 2: Significance of co-presence results.

	Q1	Q2	Q3
Kruskal Wallis H	8.712	7.653	8.106
df	3	3	3
p	0.033	0.054	0.044

In Figure 7, a graphical presentation of the results is shown, where the differences on self-reported levels of co-presence between the research conditions are more observable.

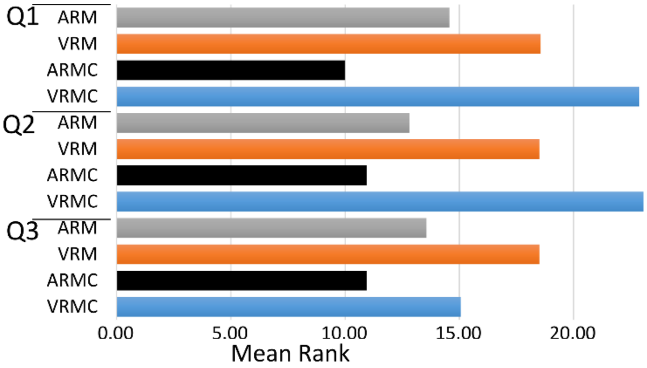


Figure 7. Graphical representation of the co-presence mean ranks question by question in each condition.

For Q1 and Q2, users wearing VR gear (VRM, VRMC) sensed more strongly that they were in the same place with the other user who was wearing AR gear (ARM, ARMC). One reason for this may be related to the used tracking technique for AR clients. As we used markers for detecting the position and orientation of the AR user, she/he was forced to look at the marker all the time or otherwise the VR user's avatar disappeared. In overall, the sense of co-presence was the strongest when the AR user's avatar was fully animated (VRMC), i.e. the AR user was wearing the motion capture suit. However, it was unexpected to see that the sense of co-presence was the lowest for AR users that were using smart glasses and the motion capture suit (ARMC). This was a phenomenon that should be further investigated.

For Q3, users wearing VR gear (VRM, VRMC) had again a stronger sense that the other user was actually present in the event. However, this time the users who saw the abstract humanoid avatar without any animations (VRM) had a stronger sense of co-presence than users who saw the fully animated human avatar (VRMC). This might be due to the fact that the more detailed human avatar did not, however, resemble with the actual person using the AR gear as the users met each other in the beginning of the user study.

We also investigated whether users' previous experience on 3D gaming, 3D virtual environments or VR HMDs had any influence on the sensed level of co-presence. The results were not statistically significant, but may indicate that previous experience on VR HMDs and 3D virtual environments could decrease the sense of co-presence in a telexistence system such as AVATAREX.

6.2 Technical Performance

For performance measurements, Qualcomm's Trepn profiler and GameBench were used. Table 3 summarizes the results for both devices in terms of (1) CPU load; (2) GPU load; (3) memory consumption; and (4) FPS. It should be noted that for Nexus 9, Trepn profiler could not measure the GPU load.

Table 3: Performance measurement results.

	CPU	GPU	Memory	FPS
ODG R7	75%	11%	1.089GB	23
Nexus 9	85%	N/A	1.775GB	15

In Table 3, the normalized CPU load is presented to improve the accuracy of the data. Normalization means that the load on the CPU is measured with respect to its maximum frequency. In our case the maximum frequency is 2.7GHz for ODG R7 and 2.3GHz for Nexus 9. It can be seen in Table 3 and in Figure 8 that both devices had a rather high average CPU load (75% and 85%). This is mainly due to the fact that the constant tracking of the marker (Kudan SDK) and the analysis of the video feed for recognizing hand gestures (Augumenta SDK) is computationally demanding. The average GPU load, on the contrary, was very low for ODG R7 (11%) because the AR client has only a few objects to render.

The average memory consumption was rather high particularly with Nexus 9. As Nexus 9 has only 2GBs of RAM, this means that on average 89% of the memory was constantly in use. Together with the high average CPU load, the lack of free memory resulted in rather low FPS (15), which could have affected the sense of co-presence when using Nexus 9. ODG R7 did not struggle with the lack of free memory, and produced higher FPS (23).

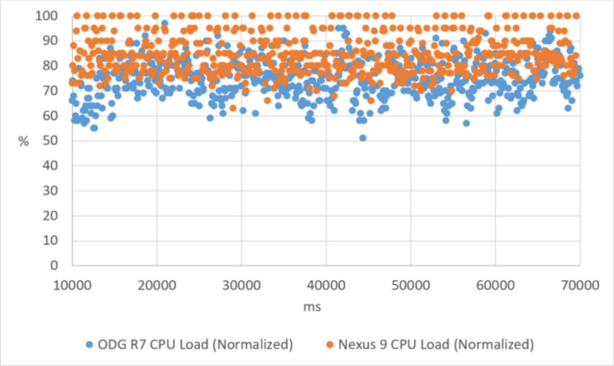


Figure 8. Average (normalized) CPU load for both devices.

7 DISCUSSION AND CONLUSIONS

In this paper, we introduced AVATAREX, a multiplatform telexistence system based on virtual avatars. AVATAREX provides means for connecting users that are simultaneously occupying the same space in the real world and its virtual replica. We investigated how people sense avatar mediated co-presence and conducted a technical evaluation with the prototype implementation of AVATAREX.

Our results show that self-reported co-presence is significantly higher when using a more immersive setup implemented with contemporary VR gear. Observing an avatar with motion capture further increased the sense of co-presence, presumably due to more realistically animated avatars. This finding is also supported by previous research [5, 7]. Users observing an avatar with AR gear sensed a lower level of co-presence when using smart glasses compared to using a mobile device. This was somewhat surprising as we expected that the smart glasses would provide more natural experience, and consequently, increase the sense of co-presence compared to use of a mobile device. There are, however, a few potential factors that may have contributed to this outcome. First, AR users may have been uncomfortable wearing the motion capture suit and/or the relatively heavy smart glasses. Second, the quality of see-through displays is still rather poor and the field-of-view is narrow (approx. 30 degrees) in the today's high-end smart glasses. In overall, this knowledge can be used in the future research and in designing telexistence systems requiring varying levels of co-presence. It should be noted that our results on avatar realism and co-presence were inconclusive and we omitted those results due to no statistical relevance. This suggests that our avatars were not quite realistic enough to observe the full spectrum of co-presence in a telexistence system such as AVATAREX.

Our prototype implementation of AVATAREX worked without any significant technical problems. Based on our performance measurements, today's high-end smart glasses (in our case ODG R7) are capable of running the prototype at a decent performance level. The average CPU load (75%) was, however, relatively high, which presumably resulted in slightly reduced FPS (23). As the smart glasses performed better than the used tablet (Nexus 9), it was surprising to see also from the technical viewpoint that the sense of co-presence was lower with the smart glasses compared to the tablet. As already mentioned, this phenomenon requires further investigation.

We acknowledge the limitations of our study. As we made a decision to change the location of the key and the chest for every participating pair, the tasks were more difficult for some participants than others. This may have affected how participants perceived co-presence. In addition, our sample size for statistical analysis was rather small which may also have affected the results, particularly concerning the questions Q4, Q5 and Q6.

In the future, we aim to study further avatar realism required to achieve a level of co-presence that is acceptable for a multiplatform telexistence system such as AVATAREX. We will also conduct user studies with a larger number of simultaneous users (already supported by AVATAREX) and extend our prototype implementation to outdoor spaces. For this, a point cloud based tracking can be used for determining the location and orientation of AR users. Furthermore, we already have access to a high fidelity 3D city model [1] that can be used as an extensive virtual replica of a real world space. Finally, we will examine how AVATAREX technology could be used in more specific use cases of remote work, for instance, in real estate management, and explore the possibilities of using other methods for assessing the full spectrum of social presence to verify the results gained from self-reported measures.

ACKNOWLEDGMENTS

The Strategic Research Council at the Academy of Finland is acknowledged for financial support of the COMBAT project (293389). This work has also been supported by the Tekes VIRPA-C project and the 6Aika: Open City Model as Open Innovation Platform project (A71143) funded by the ERDF and the City of Oulu under the Six City Strategy program.

REFERENCES

[1] Alatalo, T., Koskela, T., Pouke, M., Alavesa, P. and Ojala, T. 2016. VirtualOulu: collaborative, immersive and extensible 3D city model on the web. *21st International Conference on Web3D Technology (Web3D '16)* (2016), 95–103.

[2] Beck, S., Kunert, A., Kulik, A. and Froehlich, B. 2013. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics.* 19, 4 (Apr. 2013), 616–625.

[3] Casanueva, J. and Blake, E.H. 2001. The Effects of Avatars on Co-presence in a Collaborative Virtual Environment. *Proc Ann Conf SA Inst of Computer Scientists and Information Technologists.* (2001). DOI:https://doi.org/10.1007/978-3-7091-6785-4_10.

[4] Cummings, J.J. and Bailenson, J.N. 2016. How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology.* (2016). DOI:https://doi.org/10.1080/15213269.2015.1015740.

[5] Daher, S., Kim, K., Lee, M., Bruder, G., Schubert, R., Bailenson, J.N. and Welch, G.F. 2017. Can Social Presence Be Contagious? Effects of Social Presence Priming on Interaction with Virtual Humans. *IEEE Symposium on 3D User Interfaces* (Los Angeles, 2017), 201–202.

[6] Due, B. 2014. The future of smart glasses: An essay about challenges and possibilities with smart glasses. *Working papers on interaction and communication.* 1, 2 (2014), 1–21.

[7] Fanello, S., Rhemann, S.O.C., Dou, M., Tankovich, V., Loop, C. and Chou, P. 2016. Holoportation: Virtual 3D Teleportation in Real-time. *Chi.* (2016). DOI:https://doi.org/10.1145/2984511.2984517.

[8] Goffman, E. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings.*

[9] Harms, C. and Biocca, F. 2004. Internal Consistency and Reliability of the Networked Minds Measure of Social Presence. *Seventh Annual International Workshop: Presence 2004.* (2004).

[10] Huang, J., Quian, F., Gerber, A., Mao, Z.M., Sen, S. and Spatscheck, O. 2012. A Close Examination of Performance and Power Characteristics of 4G LTE Networks. *MobiSys.* (2012), 225–238. DOI:https://doi.org/10.1145/2307636.2307658.

[11] Huang, J., Quian, F., Guo, Y., Zhou, Y., Xu, Q., Mao, Z.M., Sen, S. and Spatscheck, O. 2013. An In-depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance. *SIGCOMM* (2013), 363–374.

[12] Häkkilä, J., Vahabpour, F., Colley, A., Väyrynen, J. and Koskela, T. 2015. Design probes study on user perceptions of a smart glasses concept. *14th International Conference on Mobile and Ubiquitous Multimedia (MUM '15)* (2015), 223–233.

[13] Jo, D., Kim, K.H. and Kim, G.J. 2015. Avatar motion adaptation for AR based 3D tele-conference. *2014 International Workshop on Collaborative Virtual Environments, 3DCVE 2014* (Jul. 2015), 29–32.

[14] Koskela, T., Pouke, M., Alavesa, P., Kukka, H., Pakanen, M., Alatalo, T. and Ojala, T. 2016. Hybrid avatars: enabling co-presence in multiple realities. *Proceedings of the 21st International Conference on Web3D Technology, Web3D 2016* (2016), 69–72.

[15] Lee, H., Ha, G., Lee, S. and Kim, S. 2017. A mixed reality tele-presence platform to exchange emotion and sensory information based on MPEG-V standard. *Proceedings - IEEE Virtual Reality* (2017).

[16] Li, Y., Snavely, N., Huttenlocher, D. and Fua, P. 2012. Worldwide Pose Estimation Using 3D Point Clouds. *Proceedings of the 12th European Conference on Computer Vision - Volume Part I.* (2012). DOI:https://doi.org/10.1007/978-3-642-33718-5_2.

[17] Lifton, J. and Paradiso, J.A. 2010. Dual reality: Merging the real and virtual. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering* (2010), 12–28.

[18] MacDonald, J.H. 2009. Kruskal-Wallis Test. *Biological Handbook of Statistics.* (2009). DOI:https://doi.org/10.1002/9780470479216.corpsy0491.

[19] Mekuria, R., Frisiello, A., Pasin, M. and Cesar, P. 2015. Network support for social 3-d immersive tele-presence with highly realistic natural and synthetic avatar users. *Proceedings of the 7th ACM International Workshop on Massively Multiuser Virtual Environments, MMVE 2015* (2015), 19–24.

[20] Milgram, P., Takemura, H., Ustimi, A. and Kishino, F. 1994. Augmented Reality: A class of display on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies.* (1994).

[21] Minsky, M. 1980. Telepresence. *Omni.* 2, 9 (1980), 45--51.

[22] Nowak, K.L. and Biocca, F. 2003. The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments.* (2003). DOI:https://doi.org/10.1162/105474603322761289.

[23] Oh, S., Park, K., Kwon, S. and So, H.-J. 2016. Designing a Multi-user Interactive Simulation Using AR Glasses. *10th International Conference on Tangible, Embedded, and Embodied Interaction (TEI '16)* (2016), 539–544.

[24] Oyekoya, O., Stone, R., Steptoe, W., Alkurdi, L., Klare, S., Peer, A.., Weyrich, T., Cohen, B., Tecchia, F. and Steed, A. 2013. Supporting interoperability and presence awareness in collaborative mixed reality environments. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST* (2013), 165–174.

[25] Poussard, B., Richir, S., Vatjus-Anttila, J., Asteriadis, S., Zarpalas, D. and Daras, P. 2014. 3DLive: A multi-modal sensing platform allowing tele-immersive sports applications. *European Signal Processing Conference* (Nov. 2014), 356–360.

[26] Slater, M. and Steed, A. 2000. A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments.* (2000). DOI:https://doi.org/10.1162/105474600566925.

[27] Slater, M. and Wilbur, S. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments.* (1997). DOI:https://doi.org/10.1007/s10750-008-9541-7.

[28] de Souza e Silva, A. 2006. Hybrid Reality Games Reframed: Potential Uses in Educational Contexts. *Games and Culture.* 1, 3 (2006), 231–251.

[29] Steed, A., Slater, M., Sadagic, A., Bullock, A. and Tromp, J. 1999. Leadership and collaboration in shared virtual environments. *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316).* (1999). DOI:https://doi.org/10.1109/VR.1999.756941.

[30] SuperData Research 2016. *Virtual Reality Industry Report 2016.*

[31] Tachi, S. 2013. From 3D to VR and further to telexistence. *23rd International Conference on Artificial Reality and Telexistence, ICAT 2013* (2013), 1–10.

[32] Tachi, S. 2016. Telexistence: Enabling Humans to Be Virtually Ubiquitous. *IEEE Computer Graphics and Applications.* 36, 1 (Jan. 2016), 8–14.

[33] Tachi, S., Tanie, K. and Komoriya, K. 1980. Evaluation Apparatus of Mobility Aids for the Blind. 1462696. 1980.