Towards a Māori Telepresence System

Jung-Woo (Noel) Park*, Steven Mills†, Hēmi Whaanga‡, Paora Mato‡, Robert W. Lindeman§, Holger Regenbrecht*
*Department of Information Science, University of Otago, Dunedin, New Zealand
parju458@student.otago.ac.nz, holger.regenbrecht@otago.ac.nz
†Department of Computer Science, University of Otago, Dunedin, New Zealand
steven.mills@otago.ac.nz
‡Te Pua Wānanga ki te Ao, University of Waikato, Hamilton, New Zealand
hemi@waikato.ac.nz, pjm20@students.waikato.ac.nz
§HIT Lab NZ, University of Canterbury, Christchurch, New Zealand
gogo@hitlabnz.org

***Abstract*—The Ātea project is investigating how technology can be used to help Māori share and preserve knowledge, culture, and language. One aspect of this is the use of telepresence systems to connect members of an iwi (tribe) or rūnanga (council) with their marae (meeting place, communal centre) and community. We describe the range of interactions we can support through voxel-based telepresence, and give examples of their application in the context of the Ātea project. We also discuss the issues around presence, as a defining element of virtual environments, that are important to this project, particularly the distinction between copresence and social presence. We show that it is possible to distinguish between these two using existing evaluation methods.**

***Index Terms*—Telepresence, copresence, virtual environments, presence measures, tikanga**

I. Introduction

Current Māori urbanisation exceeds 85 percent [1], with approximately 80 percent of tribal members living outside their tribal boundaries and 17 percent currently living in Australia [2]. With second, third and fourth generations of Māori being born in towns, cities and overseas, there is a growing concern that future generations of Māori will have little or no contact with their tribal areas and its associated knowledge systems, language, history, stories, whakapapa (genealogy), and culture. As the diaspora grows, emerging 3D telepresence technologies could help remote tribal members (re)connect with their tribal roots, but such technologies need to be deployed in a culturally appropriate manner.

The Ātea project[1] investigates technologies to preserve and share knowledge, language and culture. The name Ātea refers to a space, a wide expanse, or something that is clear and free from obstruction. It also refers to the marae ātea, an open area in front of the wharenui (meeting house), where formal welcomes to visitors take place and debates are held (Figure 1). The particular strand of this project that we discuss here is the use of virtual environments for tribal to connect back to their haukāinga (home).

In order to determine how telepresence can be used appropriately for Māori, we are developing experiences for evaluation within the Ātea project. In this paper we present the concepts of telepresence that are relevant to the project, the underpinning technologies we are developing, and the specific issues relating to Māori in this area. An initial user study is presented evaluating the effectiveness of our telepresence system, and future directions for the project outlined.

[1] https://www.sftichallenge.govt.nz/research/atea

Fig. 1. Te Rau Aroha Marae, the main building visible is the wharenui, Tahu-Pōtiki, named for the tipuna (ancestor) of the iwi. The ātea is the open space in the foreground.

II. Telepresence and Copresence

We begin by defining some terms related to concept of "presence". Although these are fundamental, the differences between them can be subtle, and they are not always used consistently in the literature. The term *presence* can be used to mean a number of things, and we draw a distinction between:

Telepresence	is a sense of presence (in any form) with a place or person who is distant from your physical location.
Spatial presence	is the sense of being in a place or environment – the sense of "being there".
Social presence	is the sense of being with another person – the sense of "being together".

Copresence	is the sense of co-location; of being in, and sharing, a place with another person – the sense of "being there, together".

Social presence and copresence are sometimes used interchangeably, but we find it useful to distinguish between them. It is possible to feel like you are "being together" without "being there" – such as when talking on the telephone to distant family. You may not, in this case, even know where they are but you can feel close to them. This is social presence, but not copresence. Similarly you can be in the same space with another person, and not feel co-present. When you pass a stranger on the street, you are both in the same place, but do not feel "together".

A. 2D/3D Telepresence Technology

Traditional videoconferencing/calling system are often used for their simplicity and portability, especially with the currently large number of mobile users. But a key disadvantage for this system is the fixed viewpoint [3], [4]. During a session users can only see what is captured by the remote camera, and therefore must always be in the camera view. This communication method is unnatural compared to face-to-face in the real world because people see their surroundings from their own viewpoint [5]. Additionally, dynamic viewpoints provide users the ability to explore an environment, which enables the sense of presence [6]. In 3D telepresence, the fixed viewpoint issue is commonly resolved using head-mounted displays (HMDs), such as the Occulus Rift, which provide users with their own egocentric view. This first-person viewpoint is important for establishing fundamental levels of presence and embodiment [6], [7].

The combination of the 3D space and egocentric, dynamic viewpoint in 3D communication systems can provide enhanced face-to-face communication compared to 2D communication systems. For example, Regenbrecht et al. [4] outline six factors that influence the social-presence experience in video-based communication systems: mutual gaze, life-sized upper body, audio/video quality, shared workspace, positioning of partners, and wide field of view. Unfortunately, we are not able to enhance eye contact with current HMD-based because they obscure the users face, but we can improve other factors for example the life-sized upper body [8]. Only displaying the upper body provides incomplete body language, but we can provide full body representations with systems such as Regenbrecht et al. and Beck et al. [9]. Representations in 3D space can help with incomplete body language, but also the relative position of partners, and shared (virtual reality) workspace.

B. Dimensions of Copresence

The human perceptual system is optimized for interactions with the real world [10], so understanding components that contribute to a 3D telepresence experience is crucial. When people communicate and interact face-to-face from body-to-body, they are typically engaging with and aware of the other person, or are copresent [8], [11], [12]. Copresence branches from two different research disciplines as Bailenson et al. [13] states. The first argues that copresence arises when users self-report that they feel like they are not alone in a virtual environment. For example, this includes research from Biocca et al. [12] who define copresence as the user's sense of being with another. The other research discipline instead argues that copresence arises when people treat embodied agents as real people [13]. In other words, if we can use human representations convincingly enough for users to believe that they co-exist in the virtual world with them, then they feel copresent. This distinction is similar to Zhao's work [14] which states that the subjective experience of being with others is different to the sociological concept of copresence. In his work, Zhao describes two dimensions of copresence: mode of copresence and the subjective feeling of copresence. The mode of copresence is related to the physical, or in our case virtual, modes of being together with others [14]. Generally, he outlines copresence in the sense of co-location, proximity, and remote interfaces, which play a critical role in achieving a subjective feeling of presence. This latter dimension is the phenomenology experience of being there with others [13] that is formed in mind of the users. Both are complementary dimensions needed when achieving human copresence because the mode of copresence describes the type of system interfaces that enable user co-location and consequently the feeling of being together with others. Through this literature, we define the concept of copresence as the mode of copresence (co-location), and social presence as the psychological phenomenon of perceiving and being with others.

C. Why 3D Telepresence in Māori Culture

Telepresence is using technology to establish a sense of shared space among remote groups and individuals [11]. Ultimately, copresence is the governing experience in achieving telepresence. However, if the mode of copresence is the main important enabler of social presence, then what type of interface is needed and how do we design one? Are simple text messaging and 2D video calling enough? Although these can support general communication among well-known friends and whānau members, they lack functionality to imbue trust and relationship building in high-communication scenarios, such as corporate business meetings [4]. Hence, there is a need for more sophisticated communication systems such as 3D telepresence. For example, body language and posture can be great visual cues to help determine what others are feeling. Especially in important formal meetings, in the absence of these cues, negotiations or trust building can easily turn sour.

In the context of Māori culture, we want to help rekindle the relationships of remote whānau and enhance the culture within younger generations. We may assume that future generations who are not born and raised within close proximity to their iwi origins would have minimal knowledge of their local history, stories, whakapapa and culture. Nourishing and connecting younger generations to their tribal identities and origins through stories and artefacts can provide a more vivid experience with 3D telepresence than a video-based communication system. Especially since these stories are generally told by local orators, which also provides the chance to be intimate and bond with storytellers and knowledgeable experts that they might have met before.

III. Voxel-based Telepresence for Ātea

In this work we use a voxel-based 3D telepresence system [15], [16] which has been shown to support the sense of presence, copresence, and embodiment. We will extend this research as part of the ongoing Ātea project.

Our telepresence system consists of two sites, each having the ability capture, record, transmit, receive, and replay 3D data streams. The capture of 3D data is through Kinect v2 sensors, although other RGB-D cameras could be used. These give a point cloud model, which is converted to a sparse voxel-based representation. The use of a voxel-based representation bounds the size of the data to be transmitted, and creates a non-photorealistic rendering enabling seamless integration of live, recorded, and virtual elements. Virtual or pre-recorded elements can be voxelised and rendered at the same level of fidelity as live transmission. Our previous work [17] has shown that this system can provide a sense of copresence, as people respond to recorded characters as if they were live, and people cannot distinguish between real and virtual (rendered) objects.

For the Ātea project, we require several different functions from the telepresence system:

- The ability for people at remote locations to share a virtual space and interact with one another.
- The ability to record people, and then play them back as virtual elements in an interactive manner.
- The ability for people (both live and pre-recorded) to interact with objects in the virtual space.

A. Supporting Collaborative Telepresence

As an example of the importance of copresence and the ability of our system to support interaction, we take the example of a hongi, a Māori greeting where two people press their noses together - literally the sharing of breath. This is an interaction that requires personal interaction between the participants, and which has great cultural significance. It links back to the creation of the first earthly woman, Hineahuone and Tāne who breathed life into her nostrils via a hongi. It is not clear whether a "virtual hongi" is possible. While similar interactions, such as handshakes, are often explored in telepresence scenarios [17], [18], sometimes with haptic elements [19], the hongi has additional cultural components. While the physical proximity can be simulated (Figure 2), in a real hongi, the two people share a breath during the hongi, and this is not something that transfers directly to the virtual space. This sharing of a breath is not just a side-effect of the interaction, but an integral part of its cultural significance.

Fig. 2. A hongi in the real world (top) and our virtual environment (bottom). While some characteristics of the hongi are captured in the virtual world, others are not.

B. Record and Replay

As well as interacting with one another, people in the virtual environment can interact with virtual or recorded avatars. The example shown in Figure 3 is a story-telling demonstrator. A kaumātua (elder) is recorded telling a story in the voxel-capture system. We identify frames where he is facing in different directions, allowing him to virtually track the motion of people in the virtual environment.

C. Object Interaction

Finally, we allow people to interact with virtual objects. The voxel-based system makes object collisions easy to detect, facilitating interactions between objects and the environment, other objects, or people (whether live or recorded). Figure 4 shows an example of an object in our virtual environment, in this case a tatā or tīrehu (bailer). The object was modelled from photographs using multi-view stereo methods, and placed in

Fig. 3. A virtual story-teller viewed from three different directions. Note that the story-teller recording is adjusted to face the viewer.

a virtual environment. The user can then pick up, move, and manipulate the object.

IV. User Study

The examples in the previous section show that we can support a range of interactions within the environment. In order to evaluate our experiences, we need to be able to assess presence. In particular, we wish to engender a sense of copresence. Given that social presence and copresence are not always clearly distinguished in the literature, we conduct an initial study to determine whether evaluation methods that aim to determine copresence measure different factors to those that aim to measure social presence.

The study was performed using a within-subjects design on 20 participants (12 female and eight male, μ_{age}=29 and σ_{age}=7.34). All participants were exposed to the same task, then copresence and social presence were measured using four well-known measurement tools, and observations were recorded. Each participant performed common social interactions first, and then did a desert-survival task.

A. Measurements

The four main "copresence" measurements come from the works of Baileson et al. [13] (BAIL), Biocca et al. [12] (BIOC), Hauber et al. [20] (HAUB), and Short et al. [21] (SHOR). Under the definitions given in Section II, BAIL and HAUB measure copresence, while BIOC and SHOR measure social presence.

Both the BIOC and HAUB measurements were evaluated on a 7-point likert scale (1-7) rated between "disagree" and "agree". The SHOR measurement was also evaluated on a 7-point likert scale (1-7), but was rated based on bipolar conditions (i.e., unsocialable/sociable and insensitive/sensitive) with no negative loadings. The BAIL measurement contains three items, with the first item having a negative loading, and each rated between "strongly disagree" and "strongly agree". The BIOC measurement contains six items with four negatively loaded items (1-2 and 5-6). The HAUB measurement has four items with the first item negatively loaded. Participants filled the questionnaires containing these 22 items based on their overall task experience in our system.

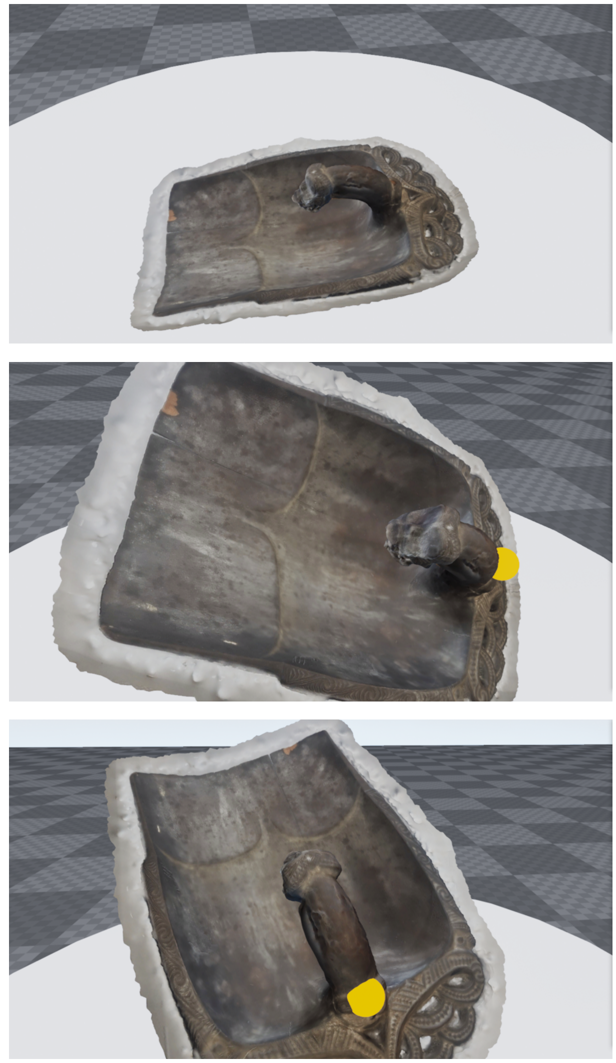

Fig. 4. Interaction with a virtual object – in this case a tatā or tīrehu (bailer) modelled using multi-view stereo techniques. The user can pick the object up, and examine it from any angle. It also interacts in a physically plausible way with other objects in the scene.

B. Procedure

When participants arrived at the room, they were given a consent and demographics form to sign and fill, and an information sheet to read, all of which took approximately three minutes. Once finished reading, we briefly explained how the system worked, the tasks they would be performing, and the measurements they need to complete at the end of the study. Then an experiment helper led one participant to a secondary room with the secondary system setup. Once the participants were separated, experimenters helped them fit their HMDs.

1) Introducing the System: Participants were introduced to the virtual environment where they would see the other participant. They were then asked four questions: Whether they could hear the experimenter through a microphone, whether they could see a white room (virtual room), whether they could see their own arms/legs, and whether they could see the person in front of them. These questions verified whether the apparatus was working correctly. During the following tasks, we recorded their behaviours.

2) Social Interaction Tasks (1 min): This small task consisted of four common social interactions to help participants get used to the system: handshake, fist bump, high five, and hugging. Participants were asked to perform these common actions in the relative order. For handshaking, fist bumping and high five actions, they were asked to perform them with both alternating hands.

3) Desert Survival Game Task (10 min): This was a modified version of the original Desert Survival Problem task from Lafferty et al. [22]. The task was simplified to accommodate a reasonable study duration. Because the task was designed to induce interaction between two people, completing the actual task was not necessary and therefore simplified.

After completing the social interaction tasks, a list of items appeared in their shared virtual space. The participants were described the following scenario:

> "In a fictional scenario, your lightweight airplane just crashed in the middle of a desert that ranges between 43-54 degress C. You two are the only survivors of the crash and have decided to stick together until rescue arrives. From the 15 items on plane, the pair of you only have time to salvage 10 items from the plane before it burns to its bare frame, in the process destroying the unsalvaged items. Please discuss and come to an agreement of the 10 items you will salvage. You have 10 minutes to complete the task."

4) Copresence Measurement and Wrap-up: When participants finished, they were asked to fill out the 22 item measurements based on their overall telepresence communication experience. After both participants finish filling their measurements, their measurements were checked for completeness and were thanked for their time. The whole procedure took less than 25 minutes.

C. Analysis

We normalised the results of the questionnaires by inverting the scale on questions where a low response was positive. The average value of each respondent on the four questionnaires was then taken as a summary of their response. Descriptive statistics for the four questionnaires are shown in Table I. All four give a mean response higher than the expected average (of 4), with BAIL having the highest overall response and HAUB the lowest. Two-tailed t-tests with $p = 0.05$ and a Bonferroni correction across the six pairs (giving an adjusted $p = 0.0083$) shows that there is a significant difference between these two, but no other pairing. While differences between the tests do not affect their reliability, this does indicate that the participants were giving varied answers.

TABLE I
SUMMARY STATISTICS OF USER RESPONSES TO SOCIAL PRESENCE AND COPRESENCE QUESTIONNAIRES.

Method	Mean	Std. Dev.
BAIL	5.8	1.03
HAUB	5.1	1.34
BIOC	5.2	1.04
SHOR	5.4	0.45

Our question is whether the copresence questionnaires (BAIL and HAUB) are measuring something different from the social presence questionnaires (BIOC and SHOR). Results of Pearson correlation co-efficient computations between the studies are shown in Table II.

TABLE II
PEARSON'S CORRELATION CO-EFFICIENTS (r-VALUES) BETWEEN THE COPRESENCE AND SOCIAL PRESENCE MEASUREMENTS EVALUATED IN OUR STUDY.

	Copresence		Social Presence	
	BAIL	HAUB	BIOC	SHOR
BAIL	1	0.625	0.454	-0.095
HAUB	0.625	1	0.357	0.199
BIOC	0.454	0.357	1	0.204
SHOR	-0.095	0.199	0.204	1

The results in Table II show that there is a correlation ($r = 0.625$) between the two co-presence questionnaires, indicating that they are measuring similar or related quantities. There are weaker correlations between the two copresence measures (BAIL and HAUB) and BIOC. Interestingly, SHOR has extremely weak correlations with the other measures. We conclude that the copresence measures appear to be consistent, and do appear to be measuring something different from the social presence measures.

V. CONCLUSIONS AND OUTLOOK

In this paper we have discussed the role of telepresence within the Ātea project. We have seen that the idea of copresense is important in this context, and discussed the various sorts of interactions that our voxel-based system supports. The concepts of social presence and copresence are not always clearly distinguished in the literature, but we have shown that questionnaires that aim to measure copresence (as we have defined it) measure different effects than those that measure social presence.

Our future work is focused on integrating the elements described in Section III to create an integrated experience based around storytelling. We will evaluate this for the strength of copresence it generates, and are interested in questions such as whether Māori have a different response, whether recordings of people can generate true copresence, and how technologies such as these can be used to connect Māori with their language, history, stories, whakapapa, and culture in a way that is respectful of tikanga, the Māori way of doing things.

Acknowledgments

This work was supported by the Science for Technological Innovation National Science Challenge project *Ātea*.

We wish to thank Ngāi Tahu, and in particular Te Rūnaka o Awarua, for their partnership and generous support and of this research.

References

[1] I. Pool, *Te Iwi Maori: a New Zealand population past, present and projected*. Auckland, New Zealand: Auckland University Press, 1991. [Online]. Available: http://www.press.auckland.ac.nz/en/browse-books/all-books/books-1991/Te-Iwi-Maori-A-New-Zealand-Population-Past-Present-and-Projected.html

[2] T. Kukutai and S. Pawar, "A Socio-demographic Profile of Maori living in Australia," 2013. [Online]. Available: https://researchcommons.waikato.ac.nz/handle/10289/7978

[3] C. Kuster, N. Ranieri, H. Zimmer, J. C. Bazin, C. Sun, T. Popa, and M. Gross, "Towards Next Generation 3D Telconferencing Systems," in *Proceedings of the 2012 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*. Zurich, Switzerland: IEEE, 2012, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/document/6365454/

[4] H. Regenbrecht and T. Langlotz, "Mutual Gaze Support in Videoconferencing Reviewed," *Communications of the Association for Information Systems (CAIS)*, vol. 37, no. 11, pp. 965 – 989, 2015. [Online]. Available: http://aisel.aisnet.org/cais/vol37/iss1/45/

[5] H. Fuchs, A. State, and J.-C. Bazin, "Immersive 3D Telepresence," *Computer*, vol. 47, no. 7, pp. 46–52, jul 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6861875/

[6] T. Schubert, F. Friedmann, and H. Regenbrecht, "The experience of presence: Factor analytic insights," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 3, pp. 266–281, jun 2001. [Online]. Available: http://www.mitpressjournals.org/doi/10.1162/105474601300343603

[7] K. Kilteni, R. Groten, and M. Slater, "The Sense of Embodiment in Virtual Reality," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 373–387, 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6797786/

[12] F. Biocca, C. Harms, and J. Gregg, "The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity," in *4th annual international workshop on presence, Philadelphia, PA*, 2001, pp. 1–9.

[8] C. Campos-Castillo, "Copresence in Virtual Environments," *Sociology Compass*, vol. 6, no. 5, pp. 425–433, may 2012. [Online]. Available: http://doi.wiley.com/10.1111/j.1751-9020.2012.00467.x

[9] S. Beck, A. A. Kunert, A. Kulik, and B. Froehlich, "Immersive Group-to-Group Telepresence," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 4, pp. 616–625, 2013. [Online]. Available: https://ieeexplore.ieee.org/document/6479190/

[10] J. Steuer, "Defining Virtual Reality: Dimensions Determining Telepresence," *Journal of Communication*, vol. 42, no. 4, pp. 73–93, dec 1992. [Online]. Available: https://academic.oup.com/joc/article-abstract/42/4/73/4210117?redirectedFrom=fulltext

[11] W. A. S. Buxton, "Telepresence: integrating shared task and person spaces," in *Proceedings of the conference on Graphics interface '92*. Vancouver, British Columbia, Canada: Morgan Kaufmann Publishers Inc., 1992, pp. 123–129. [Online]. Available: https://dl.acm.org/citation.cfm?id=155309

[13] J. N. Bailenson, K. Swinth, C. Hoyt, S. Persky, A. Dimov, and J. Blascovich, "The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 14, no. 4, pp. 379–393, aug 2005.

[14] S. Zhao, "Toward a Taxonomy of Copresence," in *Presence: Teleoperators and Virtual Environments*, vol. 12, no. 5. Cambridge, Massachusetts, USA: MIT Press, oct 2003, pp. 445–455. [Online]. Available: http://www.mitpressjournals.org/doi/10.1162/105474603322761261

[15] J.-W. N. Park and H. T. Regenbrecht, "Resolutions and Network Latencies Concerning a Voxel Telepresence Experience," *Journal of Software Engineering (JSEA)*, 2019.

[16] H. T. Regenbrecht, N. J.-W. Park, C. Ott, S. Mills, M. Cook, and T. Langlotz, "Preaching Voxels: An Alternative Approach to Mixed Reality," *Frontiers in ICT*, vol. 6, p. 7, 2019. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fict.2019.00007/full

[17] H. Regenbrecht, N. Park, C. Ott, S. Mills, M. Cook, and T. Langlotz, "Preaching voxels: An alternative approach to mixed reality," *Frontiers in ICT*, vol. 6, p. 7, 2019.

[18] J. Oh, Y. Lee, Y. Kim, T. Jin, S. Lee, and S.-H. Lee, "Hand contact between remote users through virtual avatars," in *Proceedings of the 29th International Conference on Computer Animation and Social Agents*. ACM, 2016, pp. 97–100.

[19] Z. Wang, E. Giannopoulos, M. Slater, and A. Peer, "Handshake: Realistic human-robot interaction in haptic enhanced virtual reality," *Presence*, vol. 20, no. 4, pp. 371–392, 2011.

[20] J. Hauber, H. Regenbrecht, M. Billinghurst, and A. Cockburn, "Spatiality in videoconferencing: trade-offs between efficiency and social presence," in *Proceedings of ACM CSCW'06 Conference on Computer-Supported Cooperative Work*, no. 20. Banff, Alberta, Canada: ACM, 2006, pp. 413–422. [Online]. Available: https://dl.acm.org/citation.cfm?id=1180937

[21] J. Short, E. Williams, and B. Christie, *The Social Psychology of Telecommunications*, 1st ed. New York, US: Wiley, 1976.

[22] C. Lafferty and P. M. Eady, "The Desert Survival Problem," pp. 1–9, 1974.