face2faceVR: Using AR to Assist VR in Ubiquitous Environment Usage

Yun Suen Pai
Keio University Graduate School of Media Design
4-1-1 Hiyoshi, Kohoku, Yokohama, Japan 223-8526
yspai1412@gmail.com

Hideaki Kimata
NTT Media Intelligence Laboratories, Japan
kimata.hideki@lab.ntt.co.jp

Megumi Isogai
NTT Media Intelligence Laboratories, Japan
isogai.megumi@lab.ntt.co.jp

Kai Kunze
Keio University Graduate School of Media Design
4-1-1 Hiyoshi, Kohoku, Yokohama, Japan 223-8526
kai.kunze@gmail.com

Daisuke Ochi
NTT Media Intelligence Laboratories, Japan
ochi.daisuke@lab.ntt.co.jp

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
UbiComp/ISWC'17 Adjunct , September 11-15, 2017, Maui, HI, USA
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5190-4/17/09 $15.00
https://doi.org/10.1145/3123024.3123155

Abstract
As virtual reality (VR) usage becomes more popular, one of the issues, among others, which still prevents VR from being used in a more ubiquitous manner is spatial awareness, unlike augmented reality (AR). Generally, there are two forms of such an awareness; recognizing the environment and recognizing other people around us. We propose face2faceVR; an easy to use implementation of AR tracking to assist VR towards recognizing other nearby VR users. The contribution of this work are the following; 1) it is compatible with mobile VR technology that already caters towards a wider adoption, 2) it does not require a networked or shared virtual environment, and 3) it is an inexpensive implementation without any additional peripherals or hardware.

Author Keywords
Ubiquitous VR; Virtual Reality; Augmented Reality; Spatial Awareness.

ACM Classification Keywords
H.5.1 [Information interfaces and presentation (e.g., HCI)]: Multimedia Information Systems

Introduction
VR technology today can immerse users in virtual environments never before experienced with such fidelity in the

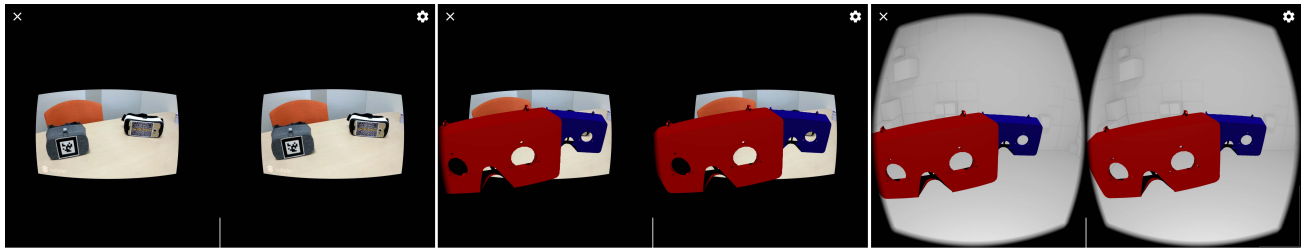

Figure 1: face2faceVR in diferent rendering stages; (left) camera being active with no rendering, (center) AR mode with superimposed virtual HMDs, and (right) background removal for VR usage.

past years. However, one of the main concern regarding VR is related to spatial awareness. VR, unlike AR, completely encapsulated the user's field of view within the head-mounted display (HMD), thus rendering the user oblivious to his or her surroundings regarding sight. This is one of the main reasons why most people believe that AR has a bigger chance of mass adoption and ubiquitous use compared to VR [7], since it augments the user's vision with virtual content through superimposition instead. Products like Google Glass have previously aimed to push AR technology toward a more public usage [2], and even though it failed to achieve their vision, it nevertheless presented an idea of these wearables replacing our smartphones that seem plausible. However, the inability for VR to become a see-through system leaves its use confined only to indoors or specifically designed environments.

In this work, we proposed face2faceVR for scenarios where multiple VR users exist within a shared physical space. In this space, each user is able to see the existence of other VR users via markers attached at the front of each HMD. In essence, we use AR's tracking technology to assist VR in terms of awareness of the surrounding VR users. This method does not require any installation on the environment, or any additional peripherals for the user to wear, except for the HMD itself. Furthermore, it was designed for mobile VR, where anyone with a recent smart phone will be able to use it. Lastly, it does not require the users to be using the same app or connected to a remote server; it merely renders the presence of other users in their own virtual space without disrupting the VR experience.

Related Work

The study on human space was coined as proxemics by Edward Halls [1], where he defined intimate, private, personal, and public space of a user depending on the proximity around him or her. Understanding human space is a key factor in VR design especially when it comes to collaborative environments or nomadic applications, where other users are present within the intimate to personal space. There have been several works by VR researchers that aim to boost user awareness during VR sessions. For example, A Dose of Reality [6] looked into altering the degree of physicality that a VR user can see by altering the rendered content. It produces a mixed reality green screen effect where the system can choose to only render the user's hands for using physical items like typing on a keyboard or rendering the full view of the users in the VR environment.

Figure 2: face2faceVR capturing a marker, running on a Google Pixel XL.

However, this method required the installation of Kinect sensors in the environment, which is possible in planned setups, but not ubiquitous use. There are also researchers that look into other forms of wearable technologies that stimulate other senses besides vision for spatial awareness, most commonly through haptics. For example, the Synesthesia Suit [3] allows for full-body haptics awareness around the user, and haptics through muscle stimulation [4]. The benefit of the haptic approach is that it caters towards those who are lacking in other senses, such as the blind [5]. However, when considering for VR, haptic devices would introduce additional peripherals over the HMD itself. Furthermore, approaches like haptic suits can greatly benefit immersion in VR, but for physical awareness, the fact that users are required to wear an entire suit for this feature in shared spaces is often deemed too troublesome.

Based on this related work, our approach can be motivated based on two scenarios; a collaborative environment where users may discuss and share the same space, and nomadic environments, where users find themselves irregularly within the space of others in varying proxemics. We are also motivated to find a solution that can be deployed on readily available mobile VR devices easily.

Implementation

We use the Unity game engine to develop the initial prototype system. The Google Daydream software development kit (SDK) was used to create the VR content, whereas the Vuforia SDK was used for AR tracking and registration. The primary hardware is the Google Pixel XL, though this is due to Daydream's compatibility at the moment. This system can be easily implemented using the Cardboard or GearVR SDK as well.

Since this is primarily a VR experience, we do not render the live feed of the phone's camera in the scene by default using background removal as shown in Figure 1, though the user can choose to toggle it at anytime using a button on the DayDream Bluetooth controller. When the rear-facing camera captures a Vuforia marker as shown in Figure 2, it renders a virtual HMD within the scene that accurately tracks the position and rotation of other VR users that are in front, thought not necessarily facing each other. Markers can be placed on the sides of the HMD too to make them visible at other angles. This allows absolute relative positioning of other HMD users to be obtained easily without relying on any sensors, making it possible to be used in public scenarios among strangers. An approach such as a networked environment would mean that the users would probably need to continuously broadcast their phone's orientation data, or be running the same application or game that is connected to a server that sends and receives the orientation data; both are unrealistic approaches. Furthermore, these methods do not pass the positional information of the phone. Using AR tracking, which is essentially computer vision, to track each user is simple and effective, limited only by the AR tracking capabilities and camera resolution.

Proposed User Study

For this work, we wish to propose a nomadic application for VR and obtain the users' feedback regarding its usability and contribution towards proxemic awareness of others. We reflect this in a study of sense of presence against awareness, as we hypothesize that rendering other users within the view would detract the user from experience, yet increases their physical awareness. The study will be carried out between the participant and multiple actors wearing the HMD in a public train environment. The training environment was chosen due to the possible randomness of a number of people present within the personal space de-

Figure 3: A collaborative environment where face2faceVR can be supported on various mobile VR platforms, since only a marker is needed.

pending on location and time of day. We plan to use three conditions; (1) baseline study where the participant is engaged in a normal VR environment, (2)semi-transparent rendering of other users within their view, and (3)solid rendering of other users within their view. We then present the participants with a questionnaire on impediment to VR usage enjoyment developed by [6].

Furthermore, we also plan to run an experiment on suitable marker designs that can be used outdoors. Currently, we are experimenting on conventional black and white markers that were specifically designed to be easily recognized, as well as Vuforia markers, which were designed to contain many features and reduce occlusion issues.

Limitations and Future Works

One of the obvious limitations with the proposed method is that this system was designed to recognize other VR users only, as shown in the collaborative environment in Figure 3. However, this work is meant to push VR towards a more ubiquitous future where VR would become more common place. Another limitation is the tracking capabilities of the Vuforia engine which handles partial occlusion well, but takes a moment to register the markers.

This current implementation makes it compatible with almost all mobile VR platforms, as it merely requires a marker to be placed on the HMD. We plan to further expand this by using Tango-enabled devices which are essential IR cameras-equipped phones that can scan both people, as well as the surrounding objects in proximity.

REFERENCES

1. Edward Twitchell Hall. 1966. The hidden dimension. (1966).
2. S Kernaghan. 2016. Google Glass: An Evaluation of Social Acceptance. (2016). https://www.researchgate.net/profile/Sheilagh_Kernaghan/publication/305340049_Google_Glass_An_Evaluation_of_Social_Acceptance/links/5789172608ae7a588ee85c56.pdf
3. Yukari Konishi, Nobuhisa Hanamitsu, Benjamin Outram, Kouta Minamizawa, Ayahiko Sato, and Tetsuya Mizuguchi. 2016. Synesthesia Suit. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 149.
4. Pedro Lopes, Sijing You, Lung-Pan Cheng, Sebastian Marwecki, and Patrick Baudisch. 2017. Providing Haptics to Walls & Heavy Objects in Virtual Reality by Means of Electrical Muscle Stimulation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1471–1482.
5. Troy L McDaniel, Sreekar Krishna, Dirk Colbry, and Sethuraman Panchanathan. 2009. Using tactile rhythm to convey interpersonal distances to individuals who are blind. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. ACM, 4669–4674.
6. Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2143–2152. DOI: http://dx.doi.org/10.1145/2702123.2702382
7. Yun Suen Pai. 2016. Physiological Signal-Driven Virtual Reality in Social Spaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 25–28.