Gestural Communication Based Pair Sightseeing System

Minghao CAI, Jiro TANAKA
Graduate School of Information, Production and Systems, WASEDA University, Japan

Abstract
In this research, we propose a gestural communication-based pair sightseeing system. It supports a gestural communication between a local user and a remote user. With the integration of Head-mounted Display and Depth Camera, we allow the local user to perform a gestural interaction with the remote user on top of the remote scene while each user is provided an independent free viewpoint. Through this system, two side of users could get a feeling that they are truly walking outdoor together side by side for a trip. We carried out a preliminary user study to evaluate our design and received positive feedback.

Keywords: Remote communication, Gestural interaction, Panoramic Viewing, Feeling together

1 Introduction

High-speed Internet and mobile communication techniques make it possible to keep in touch with someone in distance conveniently. Nonetheless, the potential of mobile video communication has yet to be fully exploited. Commercial video communication systems mostly only provide a capture of the user's face which helps little to focus on the other information like body language or the ambient or distant objects. Additionally, although might possible with current technologies, there are few communication platforms offer a way for users to achieve effective gestural communication. When users want to describe the objects or directions in the scene, only using verbal description might be challenging. Such constraints make it difficult for users to get a common perception or feel like staying together.

The problem we are targeting is helping the users in separated positions get a feeling of being together during a mobile communication. Some previous researchers have demonstrated that hand gesture is helpful in remote communication in different approaches [2, 4, 11, 12]. We find that users intend to use hand gestures to describe direction information or point out objects especially in the spatial scene, which might make the conversation smoothly.

In this work, we propose a gestural communication based virtual sightseeing system for two users in separated places: a remote user and a local user. Our research focuses on enhancing the human-to-human interaction in the mobile communication by supporting 3D air gestural communication.

2 System Design

In this system, the remote user walks around in the physical environment which would be shared, while the local user would like to have a virtual sightseeing of such shared world. The local user may have expertise related to the environment to help the remote user, or just need the surrounding to be part of the communication. We aim to realize the gestural interaction between the two users during the sightseeing. It simulates the situation that the two users walk side by side in the same physical world chatting with hand gestures. Although the two users might both stay indoors or outdoors, we assume that the local user remains indoors and the remote user goes outside in this research.

Our system's setup consists of two parts: the wearable device for the local user and the portable setup for the remote user (Figure1). Different from the traditional telepresence system, with the use of spherical camera and head-mounted display (HMD), we allow the local user to access the remote world with a 360° panoramic free viewpoint. The hand gestures of the remote user are provided directly in the capture of the remote scenery for the local user.

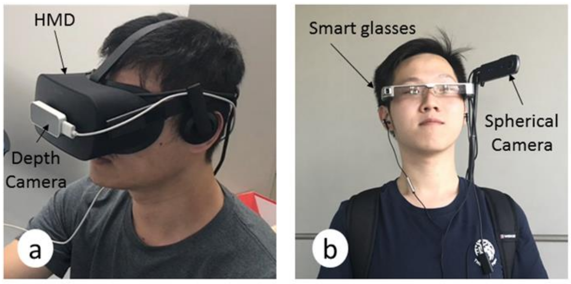

Figure 1 System setup overview

For the remote user, we introduce the augmented reality technique. By using a pair of smart glasses, our system presents the 3D air gestures of the local user directly on top of the physical world, which gives an immersive feeling.

2.1 The 360° Panoramic Browsing

This pair sightseeing system allows the local user to view the remote scenery where the remote user is.

In standard video communication like videophone call, the camera providing a remote view for the local user is carried and controlled by the remote user. In this case, the local user could not choose their own viewpoint conveniently without help from the remote one, just browsing the video more like a bystander. A certain number of different attempts have been researched to solve this restriction [1, 3, 7, 8, 9, 10]. In this work, by using a dual-fish eye spherical camera, we provide a 360° panoramic browsing of surrounding so that the local user could feel personally on the scene. Unlike the normal camera providing a limited angle of capture, our spherical camera could catch the whole 360° panoramic view in both vertical and horizontal simultaneously with no missed information.

The local user wears an HMD to see in the virtual remote scenery (Figure 2). The viewpoint is controlled by the rotation of HMD which manipulated by the local user's head movement. The local user could freely and naturally control the viewpoint by simply turning the head, just like one truly viewing in the real world, feeling personally on the scene.

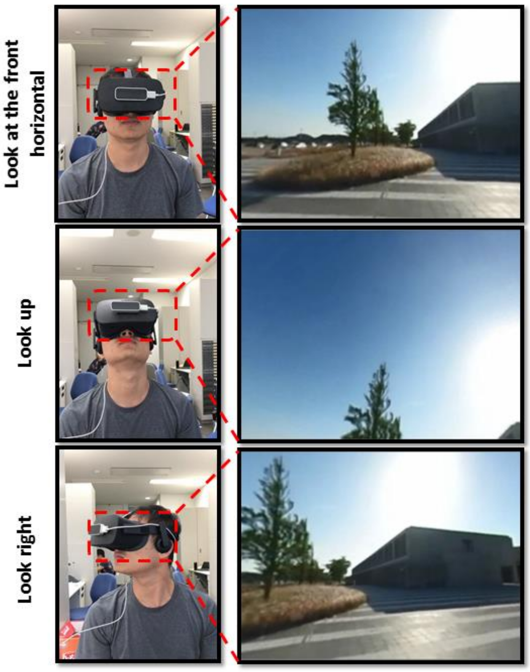

Figure 2 The viewpoint is naturally controlled by the local user's headmovement.

3 Attention Indicator

The system extracts the viewpoint data from local user's HMD and the remote user's smart glasses. By calculating the included angle between the two users' viewpoint in the remote environment, our system gives a signal to both users when they are looking at same direction (Figure 3). The system notifies the users by showing a ``SAME VIEW'' signal in the center of both users' the GUI.

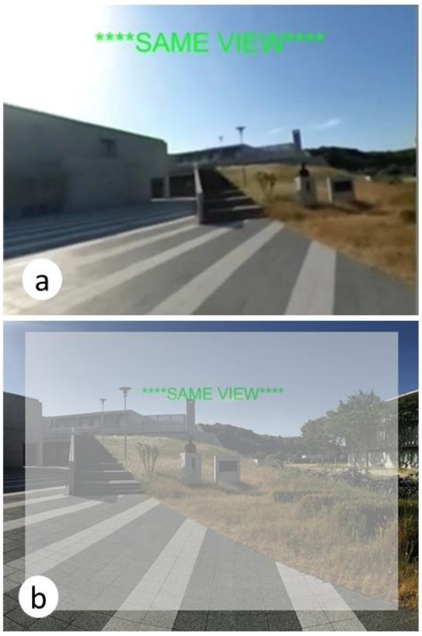

Figure 3 When the two users are viewing in the same direction, a joint attention signal would be sent to both users.

4 Gestural Communication

We choose a depth-based approach for the gesture recognition, which allows the local user completed the air gestural input freely without wearing any sensor on hands. Previous research has demonstrated that tracking the change of the depth-based bone structure could provide a high accuracy to distinguish different gestures [5, 6]. A depth camera is attached on the front side of the HMD of the local user to make sure the interactive range covering the user's viewing direction. The depth camera can extract not only the subtle changes of the spatial position and posture but also the rotation and orientation of the user's finger joints.

4.1 Gestures from the remote user

The panoramic capture of remote world for the local user includes the view of remote user's hands. The local user could directly see the hand gestures of the remote user in the remote scenery (Figure 3).

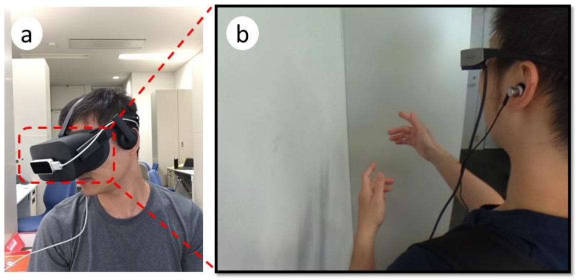

Figure 4 The local user (a) sees the remote user (b) is making gestures.

4.2 Reconstruction Hand Model

We build a pair of virtual 3D human-skin hand models to realize the gestural input of the local user. By match the hand models with the depth data of hands, the system can reappear the hand gestures of the local user in the virtual sightseeing precisely. Once the user changes the hand postures or moves the hands, the virtual models change to match the same gestures almost instantaneously.

The system presents these human-skin hand models in the local user's facing view with the First-person Perspective (FPP) on top of the remote scenery. With the use of the HMD, this design could provide an immersive virtual reality experience for the local user.

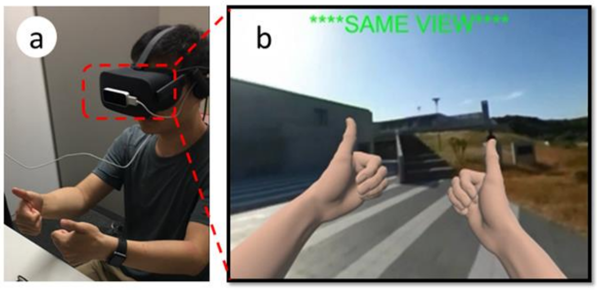

Figure 5 The local user is making gestures.

These hand models are also sent to the remote user and display on the remote user's smart glasses (Figure 6). Therefore, the remote user could see the gestures of the local user directly while viewing the environment. It is worth to point out that the remote user's perspective of the hand models is different with the local user's. The hand models are presented on the left side of the field of vision, superimposing on the physical world.

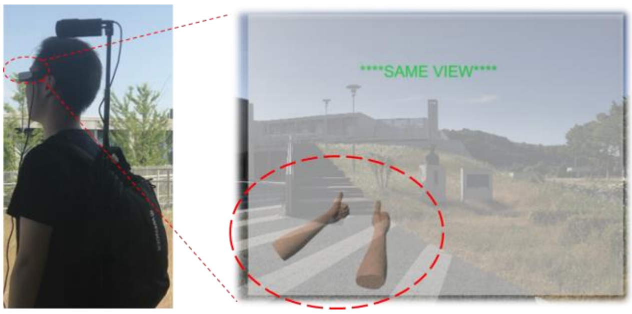

Figure 6 Visualization example of the remte user's field of view.

5 Implementation

5.1 System Hardware

Figure 6 shows the system hardware and information overview.

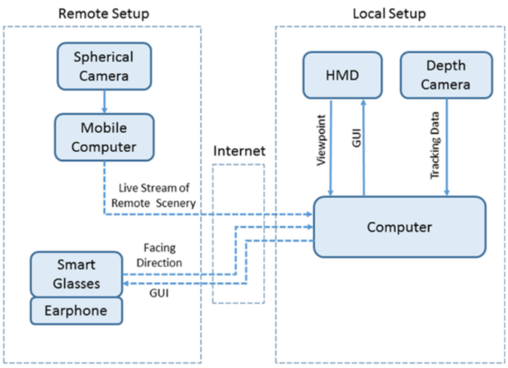

Figure 7 System hardware

5.2 Local Setup

The local setup includes the wearable devices and a desktop PC. The local user sits at the table to use our system.

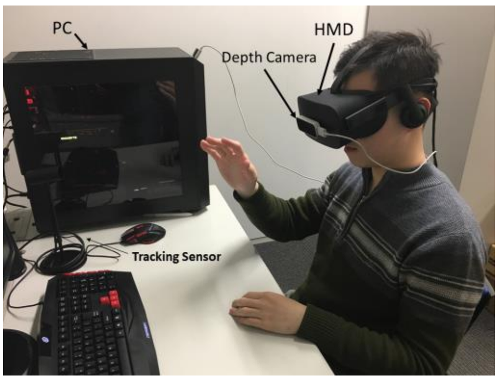

Figure 8 Local setup

The desktop PC placed on the local user side is used to analyze data and engine the core system. Unity 3D is used to render and process the incoming data from both remote and local side as well as to generate GUI for both users.

To realize the gestural recognition, we choose a new generation depth camera which has a high accuracy (an about 0.7 millimeters overall average accuracy with 8 cubic feet interactive range [13]). It is light enough (only about 45g) to make sure it is comfortable for users to wear. The effective range of the Leap Motion extends approximately from 3 to 60 centimeters above the device like an inverted pyramid.

5.3 Remote Setup

Figure 8 shows the overview of the remote setup. The remote user wears an augmented reality smart glasses which is light and compact enough (only 69 g) but supports an HD binocular displays. It packs with a motion-tracking sensor to detect the user's facing direction and a wireless module to exchange information with the local side via the internet. It presents a semitransparent display on top of the physical world while allows the user to view the physical world clearly. It provides an audio output with an earphone.

Figure 9 Remote setup

The 360° spherical camera is set on the top of a metal rod carried by the remote user. We choose this place so that the local user could see the hand gestures of the remote user (see Figure 3). The camera sends the live stream to the local user by Real Time Messaging Protocol with the help of a mobile computer.

6 Related Work

Our work is closely related to the previous research called ``WithYou", a remote communication prototype which aims to help the two users feel they go out together to some extent [1]. Comparing with WithYou, our system has some advantages in following several aspects. First, we provides an indeed 360\degree panoramic viewing for the local user while WithYou has a blind angle nearly 100° in vertical. Second, we develop a way to allow the real air gestural interaction between the two users. The users could perform gestures naturally without any wearable sensor on hands. What's more, we provide a portable augmented reality setup for the remote user, which allows the remote user to immersive in the gestures communication.

7 Conclusions

In this work, we propose our prototype system for a remote pair sightseeing between a remote user and a local user who actually far apart. By providing separated independent free viewpoint and air gestural input on top of the remote scene, we realize an intuitive air gestural communication between the two users. It simulates the local user is tripping together side by side with the remote user.

Our system is suitable for joint shopping or other possible applications like a travel guide or cooperative work.

In the future work, we plan to further improve our system. For example, we may adopt a more stable design of setup to enhance the user experience. In the future studies, we intend to implement new features that presenting an avatar of the local user in the remote scenery to enhance the user experience.

References

(1) Chang, C.-T., Takahashi, S., and Tanaka, J. A remote communication system to provide “Out Together Feeling”. Journal of Information Processing 22, 1 (2014), 76–87. Z

(2) Gauglitz, S., Nuernberger, B., Turk, M., and Höllerer, T. In touch with the remote world: Remote collaboration with augmented reality drawings and virtual navigation. In Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology, ACM (2014), 197–205.

(3) Gurevich, P., Lanir, J., Cohen, B., and Stone, R. Teleadvisor: a versatile augmented reality tool for remote assistance. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2012), 619–622.

(4) Hunter, S. E., Maes, P., Tang, A., Inkpen, K. M., and Hessey, S. M. Waazam!: supporting creative play at a distance in customized video environments. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems, ACM (2014), 1197–1206.

(5) Karam, H., and Tanaka, J. Two-handed interactive menu: An application of asymmetric bimanual gestures and depth based selection techniques. In International Conference on Human Interface and the Management of Information, Springer (2014), 187–198.

(6) Karam, H., and Tanaka, J. Finger click detection using a depth camera. Procedia Manufacturing 3 (2015), 5381–5388.

(7) Kasahara, S., and Rekimoto, J. Jackin: integrating first-person view with out-of-body vision generation for human-human augmentation. In Proceedings of the 5th Augmented Human International Conference, ACM (2014), 46.

(8) Kashiwabara, T., Osawa, H., Shinozawa, K., and Imai, M. Teroos: a wearable avatar to enhance joint activities. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2012), 2001–2004.

(9) Koizumi, S., Kanda, T., Shiomi, M., Ishiguro, H., and Hagita, N. Preliminary field trial for teleoperated communication robots. In Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on, IEEE (2006), 145–150.

(10) Ohta, S., Yukioka, T., Yamazaki, K., Yamazaki, A., Kuzuoka, H., Matsuda, H., and Shimazaki, S. Remote instruction and support using a shared-view system with head mounted display (hmd). Nihon Kyukyu Igakukai Zasshi 11, 1 (2000), 1–7.

(11) Sodhi, R. S., Jones, B. R., Forsyth, D., Bailey, B. P., and Maciocci, G. Bethere: 3d mobile collaboration with spatial input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM (2013), 179–188.

(12) Tecchia, F., Alem, L., and Huang, W. 3d helping hands: a gesture based mr system for remote collaboration. In Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry, ACM (2012), 323–328.

(13) Weichert, F., Bachmann, D., Rudak, B., and Fisseler, D. Analysis of the accuracy and robustness of the leap motion controller. Sensors 13, 5 (2013), 6380–6393.