2ZF-09

Improving the Communication of Online Open Campus with Video-Based AR and Avatar

Hao CHEN Jiro TANAKA
Waseda University

1. Introduction

Open campus is traditionally an offline event held by the school to promote the campus to the candidate students. However, due to COVID-19, online open campus has become more popular. We found that most universities use video conference software (e.g. Zoom) for online open campuses. Compared to the offline, visitors and the environment of campus are completely separated. The lack of interaction with the environment and being unable to see peers in the scene intuitively cause poor communication.

On the one hand, video is a simple method of capturing the environment, but it only has 2D information of frames. On the other hand, Augmented Reality (AR) can make the system recognize the environment in 3D and superpose virtual objects in the real world.

This paper proposes an approach of using video-based AR to capture the scenes of campus with visual information, depth information and motion information. Users from different places can watch the pre-recorded video together to have a remote group tour and use AR to communicate with others. Based on the AR capability, we use avatars to represent the remote users. By controlling the avatar, users can better communicate with other users in combination with the interaction of the scene.

2. Related Work

Priolo et al. [1] used 3D models for an interactive 3D virtual tour of the museum on web. Li, Xu, and Zhang [2] proposed a virtual reality (VR) system using 3D models and avatars for campus tours. Compared with video methods, the noise in 3D model reconstruction is more significant to worse the user experience.

Nassani et al. [3] extended the video conference with AR annotation that allows users to share their virtual pointers at a fixed depth from the screen. Kim et al. [4] proposed a remote collaboration system with virtual hand, pointer and sketch. However, some tasks are difficult to complete because the depth information of the object is not used.

3. System Design

In this system, we assume that users access the online open campus from remote places by smartphone. We will firstly introduce the technology we use for remote AR. Then introduce the sharing functions based on avatar to improve the communication.

Our approach for the online open campus is using video-based AR technology for remote AR [5].

Figure 1 shows the mechanism of the video-based AR. An AR session of general AR usually consists of the visual information from the real-time camera and motion information from sensors. While in video-based AR, the idea is packaging the intermediate data, which includes video stream, depth map stream, and the motion data, into a data set file. Then the data set file can be shared with other devices.

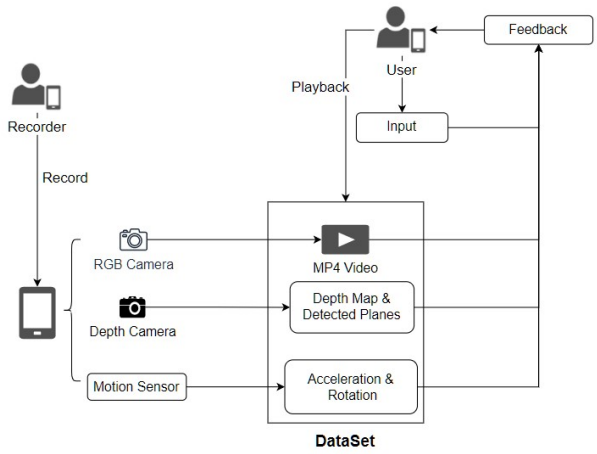


Fig. 1. Mechanism of Video-Based AR

There are two main operations in the video-based AR, recording, and playback:

Recording: This operation packages the intermediate data of the AR session into a dataset file. The general way of remote AR is to share all the data or sometimes only video data of session in real-time. While in recording, these data are correlated according to the progress of the video.

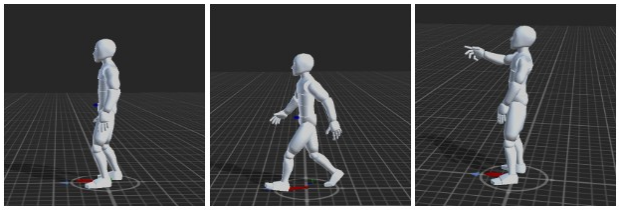
We assume the system administrator will use this function to create a data set for the campus.

Playback: As the video player restores the real image of the video according to the time axis, the 3D information and the camera pose corresponding to the current frame can be restored by using the depth map and motion data while playing the data set.

When all the users in the group are ready, the system will automatically start playing a data set from the same time to make a group tour.

Different from general remote AR, by using the intermediate data set file. The host and remote users do not need to keep real-time synchronization. Only the interaction data from group visitors need to be

transmitted. That can reduce the requirement of bandwidth. Moreover, removing the constraints of time and location make AR more convenient.



(a) Idle (b) Walking (c) Pointing

Fig. 2. Avatar Motion

Figure 2 shows the motions of avatar. In the system, the user will be represented by the avatar and see the view from the 1st-person perspective of the avatar. Also, users can see the avatar of others. Sharing functions allow the system to synchronize the actions of avatars for multi-user communication in real-time.

4. Implementation

We implemented the system with two Android phones, Google Pixel4 and Google Pixel5. And for the software, we used Unity as the platform and used the APIs of ARCore and ARFoundation.

Figure 3(a) shows the system administrator recording the scene of the campus by Android phone.

Figure 3(b) shows that two phones are placed together to simulate two remote users to playback the same data set. It can be seen from the screen that the system uses the video stream and the depth information to reconstruct the scene and the AR Session, then uses the AR annotation to visualize the plane in the video and display the avatar of another person in the peer.

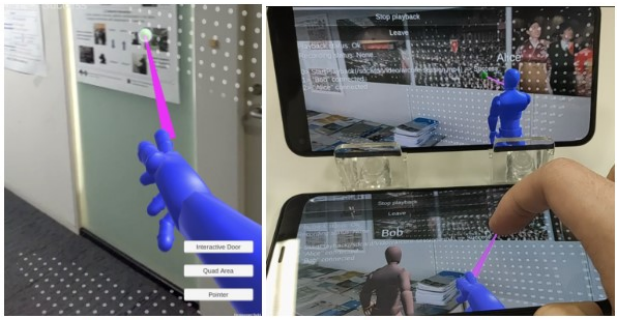
(a) Record

(b) Playback

Fig. 3. Record and Playback

Users can touch the screen to point out a place where they are interested. As Figure 4(a) shows, the pointer will be placed on the surface of physical objects based on the plane detection and depth map. Also, a virtual ray will be sent from the finger of the avatar to indicate the direction. It helps the user to interact with the real environment of campus intuitively.

In addition, we used a cloud server to synchronize the interaction of users in real-time. Figure 4(b) shows the pointing action and the pointer being synchronized between two devices.



(a) Point Out in 3D (b) Synchronization

Fig. 4. Interaction

5. Conclusion and Future Work

In this research, we discussed the problem of communication in the online open campus using video conference and the possibilities of AR prior works. Then we presented a system using video-based AR to capture the environment of the campus. It allows remote users to visit the campus and use the virtual avatar to communicate with others.

In the future, firstly we plan to conduct a user study to evaluate the improvement of communication. Secondly, we think the FoV of smartphone is limited. To improve the user experience, an idea is to integrate the 360-degree video. However, applying an external 360-degree camera to ARCore is challenging.

References

1. F. R. Priolo, R. J. B. Mediavillo, A. N. D. Austria, A. G. S. D. Angeles, M. C. G. Fernando, R. S. Cheng. (2017). Virtual heritage tour: A 3d interactive virtual tour musealisation application. ICCIP ’17. 190–195.
2. Z. Li, D. Xu, Y. Zhang. (2019). Real walking on a virtual campus: A vr-based multimedia visualization and interaction system. ICCSP ’19. 261–266.
3. A. Nassani, L. Zhang, H. Bai, M. Billinghurst. (2021). ShowMeAround: Giving Virtual Tours Using Live 360 Video. CHI EA '21. 1-4.
4. S. Kim, G. Lee, N. Sakata, M. Billinghurst. (2019). Evaluating the Combination of Visual Communication Cues for HMD-Based Mixed Reality Remote Collaboration. CHI '19. 1–13.
5. ARCore Documentation: Use AR on the go. Accessed on December, 2021. https://developers.google.com/ar/develop/unity-arf/recording-and-playback/custom-data-track