AR-based Remote Collaboration Dictation Service for Loud Environments

Eun Bee Kim, 'Daniar Estu Widiyanti, *Soo Young Shin
Department of IT Convergence Engineering,
Kumoh National Institute of Technology (KNIT), Gumi, South Korea
agrainkim@gmail.com, {'daniarestuw, *wdragon}@kumoh.ac.kr

Abstract

The AR-based remote collaboration dictation service for loud environments on the local user side is proposed. This system allows the local user to see what the remote user is saying and to send simple messages by performing the tap gesture on the virtual buttons. In the future, the live video stream from the local user and some visual cues for pointing at the components need to be added to the current application to improve the feature of the system.

Ⅰ. Introduction

Recently, Augmented Reality (AR) technology has become utilized in all areas of life. For instance, the AR technology can be used to master new skills. In AR-based learning, virtual objects are often projected in the user's head-mounted display (HMD), allowing the user to interact with the content of them [1]. To improve the user experience, multiuser AR-based learning in real-time can be helpful. With the multiple users, they can communicate each other in real-time for better understanding of the subject. The real-time cooperative work using Augmented Reality can be applied to various fields, such as engineering, education, entertainment, industrial, medicine, etc.

The Computer-Supported Cooperative Work (CSCW) is focusing on how individuals collaborate using computer technology [2]. As shown in Figure 1, four types of CSCW are presented, which are: Face-to-Face Interactions, Continuous Tasks, Remote Interactions, and Communication and Coordination. For the proposed system, the remote interactions are used.

In a general industrial remote collaboration with the AR HMDs, there are mainly two users: one is remote and the other is local. The remote user provides help to the local user, who is less experienced, to perform the maintenance or repair [3]. The local user gets guidance from the remote user such as verbal advice, interacting with shared holograms, etc. The verbal communication is critical when sharing the supplementary information which is not included in the AR application.

For the conversation between two users, usually the built-in microphones can be used. However, if the surroundings of the local AR HMD user are loud, such as in factory, the chance of the local user cannot hear what the remote user is demonstrating and vice versa is high.

In this work, the AR-based remote collaboration dictation service for loud environments on the local user side is proposed. The system aims to help each user understanding what the other user said using the help of subtitle.

Ⅱ. Application Design

Time / Space	Synchronous	Asynchronous
Co-located	Face-to-Face Interactions	Continuous Tasks
Remote	Remote Interactions	Communication and Coordination

Figure 1 Types of Computer-Supported Cooperative Work

The application design is shown as Figure 2. Two users are located in the different places, and the local user is in the loud environment.

To communicate each other using collaborative application, firstly, the remote user performs 'tap' gesture on the 'Start' button. The remote user provides the needed instruction to the local user through built-in microphone of the AR HMD, and it gets dictated by the dictation service of the collaborative application. After finishing the speech, the remote user does 'tap' on the 'Stop' button. Finally, the dictation result is shared to the local user in the text form.

For the local user, a set of buttons with the pre-made sentences are implemented so that can be sent by user performing the 'tap' gesture while gazing at the button. The local user can send some simple pre-made sentences like "yes", "no", "previous", "next", to the remote user by tapping one of the virtual buttons. Hence, the simple text is shared to the remote user in the text form.

Ⅲ. Implementation

A simple collaborative application to test the dictation service is implemented. Two Microsoft HoloLens 1 devices, one PC, and one router for networking are used. HoloLens is an Augmented Reality (AR) head-mounted display (HMD) device created by Microsoft and it is a high-performance head-mounted display.

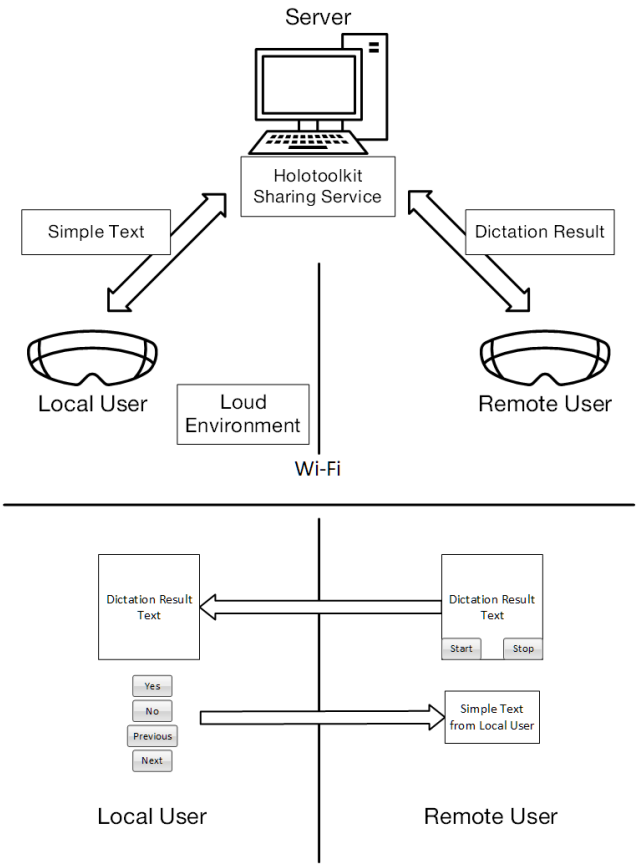

Figure 2 Application Design

Two users are wearing HoloLens in the separate places, and each user is connected to same Wi-Fi network.

For the software implementation, Unity 2018.4.26f1 version, HoloToolkit-Unity 2017.4.3.0 version to run the HTK Sharing Service on the server, Mixed Reality Toolkit-Unity (MRTK) v2.0.0, and Visual Studio 2017 to build and deploy the application to HoloLens are used. While the HTK Sharing Service is running on the server, two HoloLens are connected to the created session [4]. The HTK Sharing Service requires the IP address of the server and the specific port number "20602". In Unity, for collaboration, the InternetClient, InternetClientServer, PrivateNetworkClientServer, Microphone, SpatialPerception capabilities are enabled. And dictation feature in MRTK is used. MRTK Dictation allows users to record audio clips and acquire the result [5].

IV. Conclusion

In this paper, the AR-based remote collaboration dictation service for loud environments on the local user side is presented. To explore the proposed features, the HoloToolkit Sharing Service was used for sharing the text between the local user and the remote user. Since this work is an initialization work, in the future, the live video stream from the local user and some visual cues for pointing at the components will be added to the current application to improve the feature of the system.

ACKNOWLEDGMENT

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology"(2018R1A6A1A03024003).

References

[1] S. C.-Y. Yuen, G. Yaoyuneyong, and E. Johnson, "Augmented reality: An overview and five directions for ar in education," *Journal of Educational Technology Development and Exchange (JETDE)*, vol. 4, no. 1, p. 11, 2011.

[2] J. Grudin, "Computer-supported cooperative work: history and focus," *Computer*, vol. 27, no. 5, pp. 19–26, 1994.

[3] R. A. J. de Belen, H. Nguyen, D. Filonik, D. Del Favero, and T. Bednarz, "A systematic review of the current state of collaborative mixed reality technologies: 2013–2018," *AIMS Electronics and Electrical Engineering*, vol. 3, no. 2, pp. 181–223, 2019.

[4] keveleigh, "Hololens (1st gen) sharing 240 - multiple hololens devices - mixed reality." [Online]. Available: https://docs.microsoft.com/en-us/windows/mixed-reality/develop/unity/tutorials/holograms-240

[5] "Dictation — Mixed Reality Toolkit Documentation." [Online]. Available: https://microsoft.github.io/MixedRealityToolkit-Unity/version/releases/2.0.0/Documentation/Input/Dictation.html