Regular Paper

Proposal of a Broadcaster Support Method using MR Stamp in 360-degree Internet Live Broadcasting

Yoshia Saito* and Kei Sato*

*Graduate School of Software and Information Science, Iwate Prefectural University, Japan
y-saito@iwate-pu.ac.jp

Abstract - In this study, we investigate the use of Mixed Reality (MR) stamp which supports broadcasters in reducing communication errors from viewers to a broadcaster in 360-degree internet live broadcasting. There is a problem that the broadcaster is unable to grasp the viewer's POV (Point Of View) compared to the conventional broadcasting method. We have confirmed that this problem could be reduced by combining an equirectangular video and 2D stamp which presents the viewer's interests in a simple image on the video. On the other hand, the 2D stamp system has three issues. The issues are (1) the stamp cannot be fixed on the target object, (2) the broadcaster must check a PC display to see the stamp, and (3) it is difficult to understand the position of the stamp in real space.
In this study, we propose MR stamp which can be displayed and fixed on the real space, which enables the broadcaster to check a holographic stamp on the real space through an MR device. To realize our proposal, we implemented the proposed system using Microsoft's HoloLens 2 which is a head mounted display (HMD) as the MR device. It can show holograms in real space by recognizing real space. We also evaluated the effectiveness of the MR stamp compared with the 2D stamp and found that the MR stamp with the spatial audio solved the three issues of the 2D stamp.

Keywords: 360-degree internet live broadcasting, Mixed Reality, MR stamp, Broadcaster support method

1 INTRODUCTION

YouTube started a 360-degree internet live broadcasting service from 2016 and it enables anyone to easily use the 360-degree internet live broadcasting service now. 360-degree internet live broadcasting is a service that combines internet live broadcasting with 360-degree videos using an omnidirectional camera. In 360-degree internet live broadcasting, a broadcaster can provide a 360-degree video to viewers in real-time without caring about the view angle of the camera. The viewers can change the POV according to their interests and communicate with the broadcaster using text chat.

The 360-degree internet broadcasting, however, has a lot of new problems. One of the problems is that the broadcaster cannot be aware of the viewers' POV. In the conventional internet live broadcasting, it uses a web camera which has a single lens and the single lens definitely shows the viewers' POV. The broadcaster can see what they are watching by the direction of the lens. On the other hand, 360-degree internet live broadcasting uses an omnidirectional camera that has a wide-angle lens or multiple lenses. It prevents the broadcaster from seeing what the viewers are watching by the direction of the lens.

There are many studies about the role of gaze information in remote communication [1][2]. In the studies, it concludes that the communicatee's gaze information indicates the target of interest or center of the topic. The gaze information in the remote communication is similar to the viewers' POV in the 360-degree internet live broadcasting. The broadcaster sometimes cannot understand the context of the viewers' comments and it causes communication errors between the broadcaster and the viewers.

To solve the problem of communication errors, we have proposed stamp functions in 360-degree internet live broadcasting [3]. In this study, the stamp functions help the broadcaster to understand what the viewers are talking about and find the object. However, it remains three issues in the proposed system, which are (1) the stamp cannot be fixed on the target object, (2) the broadcaster must check a PC display to see the stamp, and (3) it is difficult to understand a position of the stamp in the real space. These issues should be solved to realize smooth communication between the broadcaster and the viewers.

In this study, we propose MR stamp which is a new stamp function using MR that can be displayed and fixed in real space. It enables the broadcaster to check a holographic stamp on the real space through an MR device. The MR stamp can solve the issues of the previous stamp function.

The contributions of this paper are summarized as follows:

- We proposed a new broadcaster support method called MR stamp in 360-degree internet live broadcasting.
- We developed and evaluated a prototype system of the MR stamp using HoloLens 2.
- We clarified that the MR stamp enabled the broadcaster to find a target object in a relatively short time and spatial audio could help the broadcaster to find the MR stamp.

The rest of this paper is organized as follows. Section 2 describes our previous work about stamp functions in 360-degree internet live broadcasting. Section 3 describes an overview of the proposed system and a use case of the MR stamp. Section 4 describes the implementation of the proposed system. Section 5 describes a first evaluation experiment to clarify the effects of the proposed system and reveal its problem. Section 6 describes a second evaluation experiment improving the proposed system. Section 7 summarizes this study.

2 PREVIOUS STUDY

We have proposed stamp functions in 360-degree internet live broadcasting to support communication between the broadcaster and the viewers. In this section, we explain the stamp functions in the previous study and its known issues.

2.1 Stamp Functions in 360-degree Internet Live Broadcasting

There are two stamps which are “Look” and “Go” stamps in the previous study as shown in Fig. 1. Figure 2 shows a user interface of the viewer in the previous study. The viewers can watch the 360-degree live video in spherical format and change the POV as they want. The viewers are also able to use the Look and Go stamp by selecting the kind of stamp and clicking on the video. The stamps are shown in the same place as the video on the user interface of the viewers and the broadcaster. Figure 3 shows the user interface of the broadcaster. The 360-degree video is displayed in equirectangular format. The broadcaster can check the stamps that are sent from the viewers without changing the POV.

We evaluated the previous stamp function and found their advantages. The stamp function improved the easiness of communication between the broadcaster and the viewers compared with the case that the stamp was not used. The broadcaster could easily understand what the viewer talked about and found the object. Moreover, the stamp function increased the frequency of communication between the broadcaster and the viewers.

2.2 Known Issues

Although the previous study has advantages, there are three issues as follows.

(1) The stamp cannot be fixed on the target object.

A stamp only has information of a direction from the omnidirectional camera at a particular time. Therefore, the stamp cannot be fixed on the target object and it causes a positional shift of the stamp when the omnidirectional camera is moved.

(2) The broadcaster must check a PC display to see the stamp.

Figure 1: Stamps in the previous study.

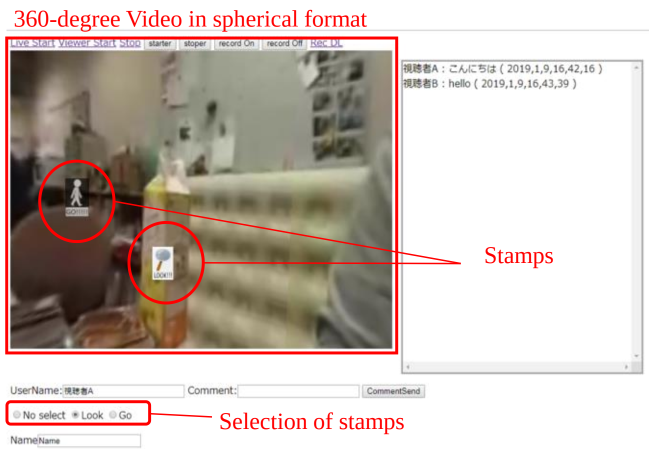

Figure 2: User Interface of viewer in the previous study.

Figure 3: User Interface of broadcaster in the previous study.

In the previous study, the broadcaster has to carry a laptop PC and check stamps from the viewers through the display. The check of the laptop PC display consumes time and prevents smooth communication between the broadcaster and the viewers. Besides that, it is dangerous to see the laptop PC while walking and it is inconvenient not to be able to use a hand which keeps the laptop PC.

(3) It is difficult to understand the position of the stamp in the real space.

The stamps are shown on the 360-degree video in equirectangular format. When a stamp appears on the video, the broadcaster must find the position of the stamp in real space. The distortion of the 360-degree video in equirectangular format makes it difficult to find the direction of the stamp.

3 PROPOSED SYSTEM

The MR stamp aims to solve the issues of the previous study using MR and realize smooth communication between the broadcaster and the viewers.

3.1 Effectiveness of MR

MR is a technology that displays holograms of virtual objects in real space and the users can interact with the holograms. Several researches show the effectiveness of MR in remote communication between users.

Lee [4] developed an MR remote collaboration system that shared 360-degree live video. In this system, a hologram of the remote user's hand is displayed in real space through the MR device. The hand gestures by the hologram help to understand each other's focus and improve their communication. Johnson [5] studied the effect of MR guidance. An experiment was conducted to understand how to provide explicit spatial information in a collaborative MR environment. The experiment result showed the MR guidance realized effective referencing through deixis. Other several researches show the effectiveness of the hologram for remote communication [6, 7]

From the related work, the reduction effect of communication errors can be also expected by introducing the MR technology for the stamp function in 360-degree internet live broadcasting.

3.2 System Model

Figure 4 shows a model of the proposed system. The proposed system is included in the existing 360-degree internet live broadcasting system. A broadcaster provides 360-degree live video to viewers using the 360-degree internet live broadcasting system. The viewers can send a 2D stamp which is implemented in the previous study to the broadcaster. The proposed system receives the 2D position information of the stamp and transforms it into 3D position information. The proposed system displays an MR stamp using the 3D position information and the broadcaster can check the MR stamp in real space through an MR device.

The proposed system can solve three issues in the previous section. The first issue which is that "The stamp cannot be fixed on the target object" can be solved by fixing the stamp on the real space using MR. The second issue which is that "The broadcaster must check a PC display to see the stamp" can be solved by using an MR device which is a type of HMD. The third issue which is that "It is difficult to understand the position of the stamp in the real space" can be solved by directly displaying the stamp on the target object in the real space using MR.

Table 1: Comparison between 2D stamp and MR stamp.

	Time to find	Errors to find	Easiness to find
2D stamp	Long time	Many errors	Difficult
MR stamp	Short time	Fewer errors	Easy

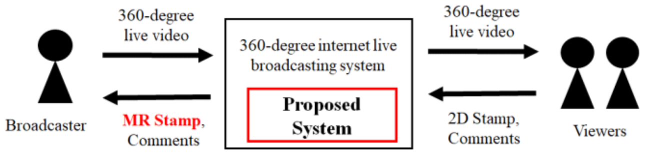

Figure 4: A model of the proposed system.

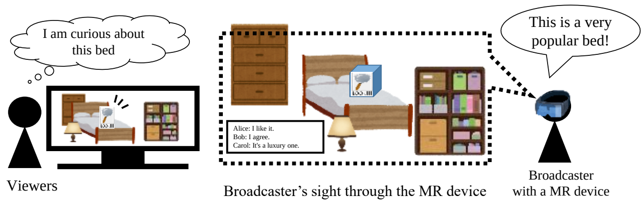

Figure 5: A use case of the proposed system.

Table 1 shows a comparison between the 2D stamp of the previous study [3] and the MR stamp. In terms of time to find the target object, the MR stamp would shorten the time because it does not need to check a PC display and understand the position of the stamp in the real space. In terms of the number of communication errors to find the target object, the MR stamp would also reduce the errors because it can be fixed on the target object. Moreover, the broadcaster would find the MR stamp easier to find the target object than 2D stamps because it has the advantage of less time and errors to find the target object besides that the MR device frees a hand which keeps the laptop PC.

3.3 Use Case

In this study, we suppose that an omnidirectional camera is fixed on an arbitrary position and 360-degree internet live broadcasting is performed indoors. The reason is that the usage environment should be simple for the first step of this study. We also suppose it is used at the showroom and the exhibition.

Figure 5 shows a use case of the proposed system. The viewers can send an MR stamp to the broadcaster if they want to see a particular object in a showroom. The broadcaster can understand the request from the viewers easier than when only comments are used. Since 2020, online virtual events have been increasing because of COVID-19. In online virtual events, it is said that more opportunity for real two-way communication between the broadcaster and viewers [8]. In online conferences, it is reported that they cannot have smooth relationships with each other between participants [9]. The MR stamp would realize smooth two-way communication between the broadcaster and the viewers in various online virtual events.

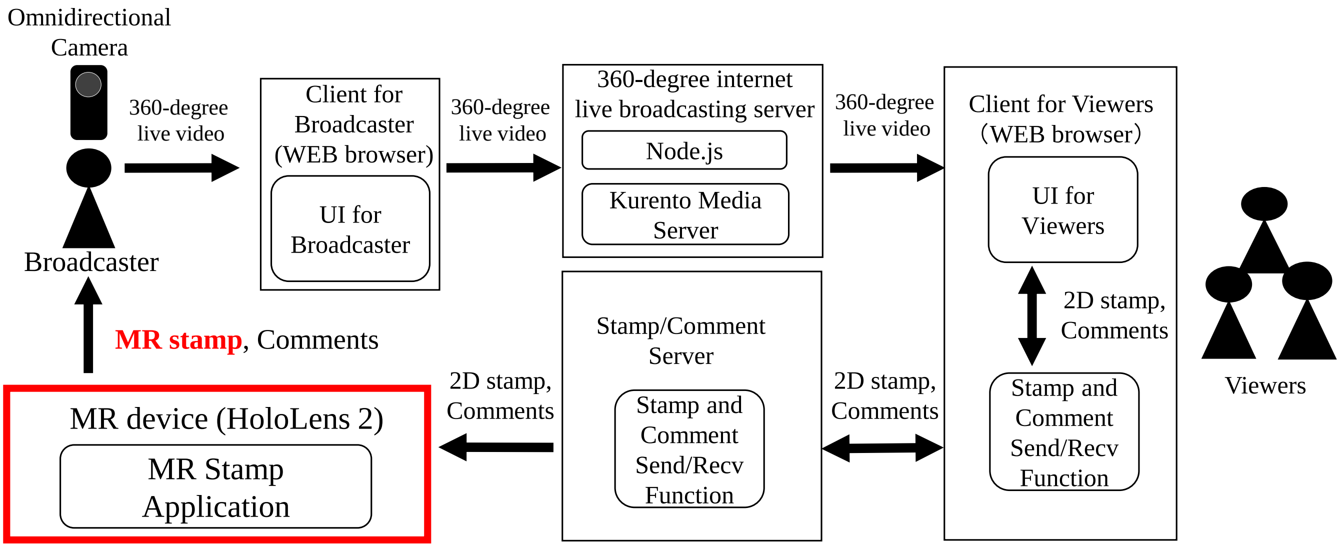

Figure 6: Architecture of the prototype system.

4 IMPLEMENTATION

We implemented a prototype system of the MR stamp using Hololens2. In this section, we describe the architecture of the prototype system and its main application.

4.1 System Architecture

The prototype of the proposed system is based on the 360-degree internet live broadcasting system of the previous study. Figure 6 shows the architecture of the prototype system. The red square shows new implementation in this study and the other parts are diverted from the previous study. A broadcaster can start 360-degree internet live broadcasting using the client for broadcaster on a web browser. The 360-degree internet live broadcasting server distributes it to viewers. The viewers can watch the 360-degree live video and send 2D stamp and comments in the same manner as the previous study. The stamp/comment server forwards it to all clients for viewers and an MR device of the broadcaster. We use Microsoft HoloLens 2 as the MR device. In the HoloLens2, the MR stamp application is running. The application presents the MR stamp and comments to the broadcaster. The broadcaster can check the MR stamp in real space through the HoloLens 2.

4.2 MR Stamp Application

The MR stamp application receives 2D position information of the stamp and needs to transform it to 3D position information for the MR stamp. To realize the coordinate transformation, we use a Raycast function in Unity, which irradiates a 3D ray from an origin point to a target direction and detects intersecting collisions. The origin point is the coordinates of the omnidirectional camera. The target direction can be given by the coordinates of the stamp. The HoloLens 2 has a function of spatial mapping. The spatial mapping provides a detailed representation of real space surfaces in the environment around HoloLens 2. The raycast detects the intersecting collisions with the real space surfaces and returns the 3D coordinates. The MR stamp is displayed on the 3D coordinates.

Figure 7 shows an example of the MR stamp. The MR stamp is shown as a 3D square frame so that it is easy to check the target object from the broadcaster. The MR stamp disappears after a period of time (10 seconds in this implementation).

The comments from the viewers are displayed on a comment window as shown in Fig. 8. The comment window tracks the broadcaster's sight.

Figure 7: An example of the MR stamp.

Figure 8: Comment window through HoloLens 2.

5 FIRST EVALUATION

We conducted an experiment to evaluate the effectiveness of the MR stamp using the implemented prototype system. The purpose of the evaluation is to check whether the three issues of the previous study can be solved or not by using the MR stamp.

5.1 Environment and Procedure

We compare the prototype system (MR stamp system) with the stamp system of the previous study (2D stamp system). In the experiment, A broadcaster performs 360-degree internet live broadcasting and two viewers watch the broadcasting. In the room of the broadcaster, there are various objects. The viewers talk about an object in the broadcaster's room using the MR/2D stamp and comments. The broadcaster looks for the target object and communicates with the viewers. The evaluation items are as follows; (1) time to find the target object, (2) number of communication errors, and (3) subjective easiness to find the target object.

The experiments were conducted 4 times. There were one broadcaster and two viewers per time and the participants were students at our university. The broadcasting time was 30 minutes. Figure 9 shows the procedure of the experiment. At first, a viewer sends a stamp to the broadcaster. The target object and the viewer who performs the task are predetermined by the task instruction. The broadcaster looks for the target object referring to the stamp. After the target object is found, the broadcaster confirms whether the object is correct or not by speaking to the viewer. The viewer replies whether it is correct or not. If it is not correct, the broadcaster continues to find the target object. The search task with stamp is performed 2 times. At last, we ask the broadcaster how easy to find the target object in 5-point scale with a questionnaire. This procedure was performed for both the MR stamp system and the 2D stamp system. The order of the MR stamp and 2D stamp system was random to keep fairness.

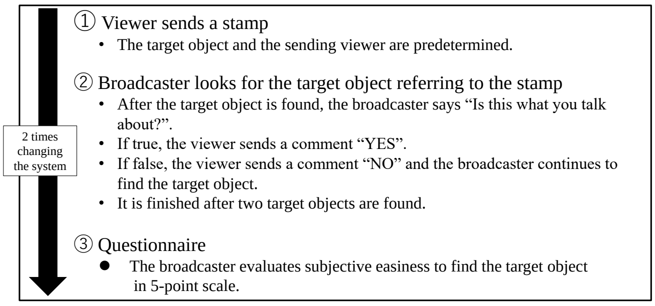

Figure 9: Procedure of the experiment.

Figure 10: Target objects.

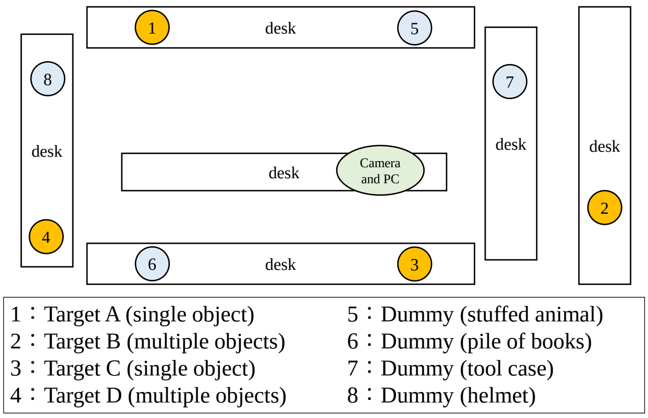

Figure 11: Location of all objects in the broadcaster's room.

Figure 10 shows the target objects used in the experiment. Target A and C are comprehensible ones to find because there is nothing around the object (hereafter, single object). Target B and D are mistakable ones to find because there are several similar objects around the object (hereafter, multiple objects). In addition to these target objects, there are several dummy objects in the broadcaster's room. We locate the target and dummy objects scattered in all directions to make the broadcaster look around to find the object. Figure 11 shows the location of all objects in the broadcaster's room.

5.2 First Evaluation Results

Figures 12 and 13 show the evaluation results of the first evaluation. There are the results of 4 broadcasters with 2D and MR stamp systems. In terms of Target, "Single" means the single object such as Target A and C, and "Multiple" means multiple objects such as Target B and D. "Time to find" denotes the time from when a stamp was displayed to when the broadcaster found the correct target object. "Errors" denotes the number of times when the broadcaster indicated incorrect objects.

The average time to find the target object in the 2D stamp system is 11.25 seconds in the case of the single object and 20.25 seconds in case of the multiple objects. The total of the average time in the 2D stamp system is 15.75 seconds. In the MR stamp system, it takes 18 seconds in the case of the single object and 13.25 seconds in case of the multiple objects. The total of the average time in the MR stamp

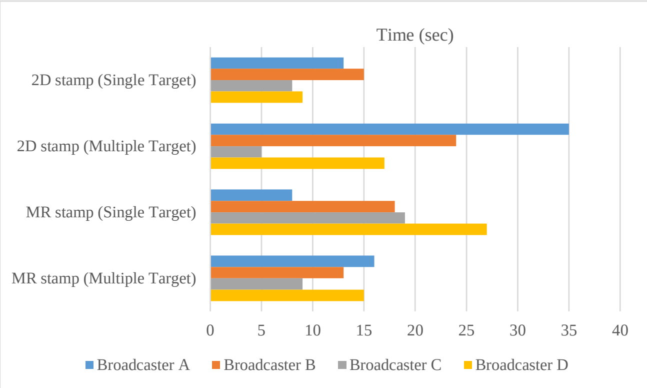

Figure 12: Evaluation result of time to find the target object in the first evaluation.

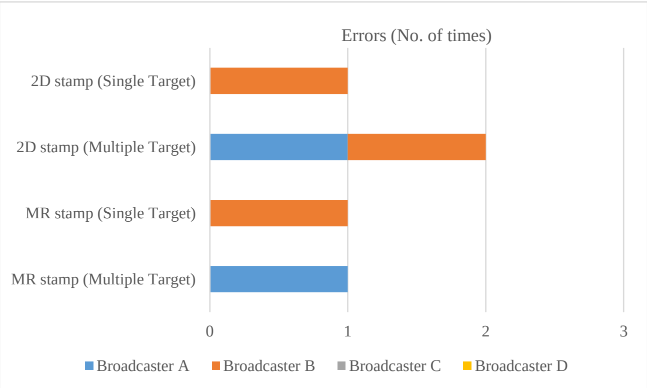

Figure 13: Evaluation result of the number of errors to find the target object in the first evaluation.

system is 15.625 seconds. The average number of errors in the 2D stamp system shows 0.25 and 0.5 in the single and multiple objects respectively. The total of the average number of errors in the 2D stamp system is 0.375. The average number of errors in the MR stamp system shows 0.25 and 0.25 in the single and multiple objects respectively. The total of the average number of errors in the MR stamp system is 0.375. We conducted a Student's t-test for the results of the total average and there is no significant difference in the average time to find ($p = 0.98$) and the average number of errors ($p = 0.63$) between the 2D and MR stamp systems.

Figure 14 shows the result of the questionnaire which is about the subjective easiness to find the target object in 5-point scale. The average points of the 2D stamp system are 4.5 and 3.25 in the single and multiple objects respectively. The total average in the 2D stamp system is 3.85. The average points of the MR stamp system are 4.25 and 3.75 in the single and multiple objects respectively. The total average in the MR stamp system is 4. We conducted a Student's t-test for the results of the total average and there is no significant difference in the subjective easiness to find the target object ($p = 0.83$).

From these results, we didn't find the effectiveness of the MR stamp system. In the free descriptive answer of the questionnaire, there were several same answers that "It was difficult to find the MR stamp". Even if the MR stamp can help the broadcaster find the target object, the broadcaster

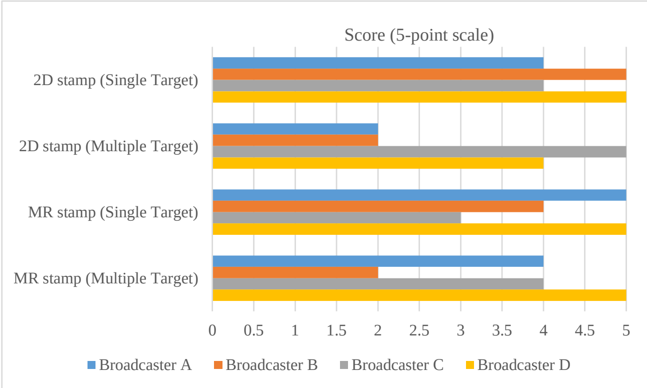

Figure 14: Questionnaire result of subjective easiness to find the target object in the first evaluation.

cannot find the location of the MR stamp itself in the experiment. We expect that an additional function to find the location of the MR stamp will improve the effectiveness of the MR stamp system.

6 SECOND EVALUATION WITH SYSTEM IMPROVEMENT

In the first evaluation, there was no difference in the effectiveness between the 2D stamp system and the MR stamp system. The viewing angle of HoloLens2 is approximately 52 degrees diagonal (28.5 degrees vertical, 43 degrees horizontal), which is narrower than the human viewing angle. Because it can only see MR stamps within a limited range. it is difficult to find the location of the MR stamps within a certain time. Since the prototype system had a problem in that it was difficult to find the location of the MR stamp itself, we tried to improve the prototype system to find the MR stamp easily and conduct a second evaluation using the improved prototype system.

6.1 System Improvement

To help the broadcaster find the MR stamp, we introduce spatial audio into the MR stamp. The spatial audio enables the broadcaster to perceive sound all around and be aware of the location of the sound source. Titus [10] studies the effectiveness of spatial audio in finding the location of a maker. He finds that the spatial audio helps the user with an HMD to find a rough location of the maker although it is not enough to specify the exact location of the maker. In our MR stamp system, the broadcaster would find the MR stamp if he/she can be aware of a rough location of the stamp.

We implemented the spatial audio to the MR stamp using a spatial audio function of MRTK (Mixed Reality Toolkit). When the MR stamp is displayed, the spatial sound is generated from the location of the MR stamp. The broadcaster can hear the spatial sound through the HoloLens 2 and find the MR stamp quickly.

We also conducted a preliminary experiment to check how accurately the user could grasp the direction of the spatial sound using HoloLens 2. From the experiment, the user could grasp the horizontal direction of the spatial sound approximately. Although the accuracy was low and it was not sufficient to grasp the accurate direction, it could support the MR stamp. On the other hand, it was hard to grasp the vertical direction of the spatial sound in this implementation. Since the target objects were placed horizontally in the first evaluation, the problem that the vertical direction could not be grasped would not be affected.

6.2 Second Evaluation Results

We conducted a second evaluation using the improved prototype system with spatial audio. The environment and the procedure are the same as the first evaluation.

Figures 15 and 16 show the evaluation results of the second evaluation. The average time to find the target object in the 2D stamp system is 17.5 seconds in the case of the single object and 19.75 seconds in case of the multiple objects. The total of the average time in the 2D stamp system is 18.625 seconds. In the MR stamp system, it takes 5 seconds in the case of the single object and 5 seconds in case of the multiple objects. The total of the average time in the MR stamp system is about 5 seconds. The average number of errors in the 2D stamp system shows 0.25 and 0.5 in the single and multiple objects respectively. The total of the average number of errors in the 2D stamp system is 0.375. The average number of errors in the MR stamp system shows 0 and 0 in the single and multiple objects respectively. The total of the average number of errors in the MR stamp system is 0. We conducted a Student's t-test for the results of the total average. There is a significant difference in the average time to find ($p = 0.01$) and a significant trend in the average number of errors ($p = 0.08$) between the 2D and MR stamp systems. Although the errors in the 2D stamp would be caused by the first issue "The stamp cannot be fixed on the target object", no errors occurred in the MR stamps.

Figure 17 shows the result of the questionnaire which is about the subjective easiness to find the target object in 5-point scale. The average points of the 2D stamp system are 3 and 2.25 in the single and multiple objects respectively. The total average in the 2D stamp system is 2.625. The average points of the MR stamp system are 4.75 and 4.5 in the single and multiple objects respectively. The total average in the MR stamp system is 4.625. We conducted a Student's t-test for the results of the total average and there is a significant difference in the subjective easiness to find the target object ($p = 0.0007$).

From these results, we can find the effectiveness of the MR stamp system with special audio. In terms of the three issues of the previous study, the first issue which is that "The stamp cannot be fixed on the target object" is solved because the number of errors decreases in the proposed system. The second issue which is that "The broadcaster must check a PC display to see the stamp" is solved because the time to find the target object decreases. The third issue which is that "It is difficult to understand the position of the stamp in the real space" is solved because both the number of errors and the time to find the target object decreases.

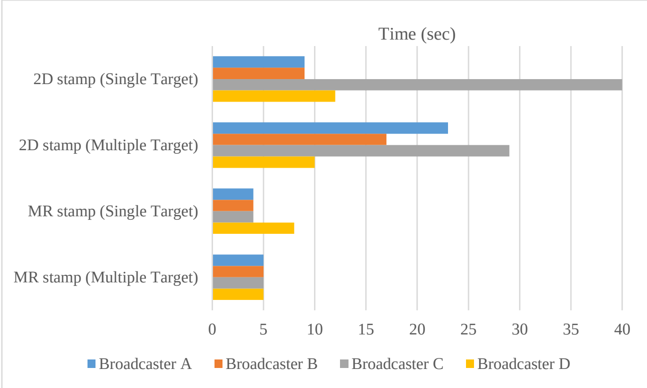

Figure 15: Evaluation result of time to find the target object in the second evaluation.

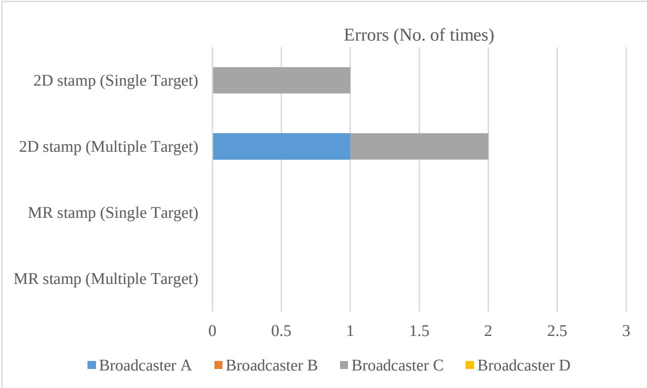

Figure 16: Evaluation result of the number of errors to find the target object in the second evaluation.

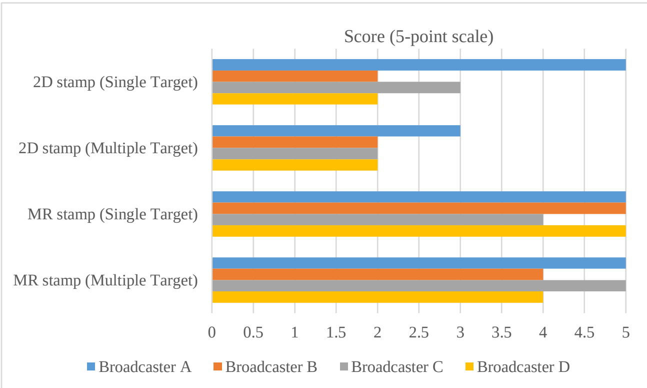

Figure 17: Questionnaire result of subjective easiness to find the target object in the second evaluation.

7 CONCLUSION

In this study, we proposed MR stamp which can be displayed and fixed on the real space, which enables the broadcaster to check a holographic stamp on the real space through an MR device. We implemented a prototype system of the MR stamp with special audio. In the evaluation of this study, we verified the effectiveness of the MR stamp compared with the 2D stamp using the prototype system.

We found that the MR stamp with the spatial audio made it easier for the broadcaster to intuitively grasp the location of the stamp. In addition, a comparison with the 2D stamp system revealed that there was a significant difference in the time to find the target object and the subjective easiness to find the target object. There was also a significant trend in the average number of errors between the 2D and MR stamp systems. From the result, we found that the MR stamp could solve the issues of the 2D stamp. In the future, it is necessary to improve the accuracy of the stamp display.

REFERENCES

[1] R. Vertegaal, “The GAZE groupware system: mediating joint attention in multiparty communication and collaboration”, Proc. of SIGCHI conference on Human Factors in Computing Systems, pp.294-301 (1999).
[2] D. M. Grayson and A. F. Monk, “Are you looking at me? Eye contact and desktop video conferencing”, ACM Transactions on Computer-Human Interaction (TOCHI), Vol. 10, Issue 3, pp.221-243 (2003).
[3] Y. Saito, A. Kuzumaki, Y. Yahata and D. Nishioka, “Development of a Communication Support System with Stamp Functions in 360-degree Internet Live Broadcasting”, Proc. of DICOMO2019, pp.895-900 (2019) (in Japanese).
[4] G. A. Lee, T. Teo, S. Kim and M. Billinghurst, “Mixed Reality Collaboration through Sharing a Live Panorama”, Proc. of SA’17, pp.1-4 (2017).
[5] J. G. Johnson, D. Gasques, T. Sharkey, E. Schmitz and N. Weibel, “Do You Really Need to Know Where “That” Is? Enhancing Support for Referencing in Collaborative Mixed Reality Environments”, Proc. of SIGCHI conference on Human Factors in Computing Systems, pp.1-14 (2021).
[6] O. Oda, C. Elvezio, M. Sukan, S. Feiner and B. Tversky, “Virtual Replicas for Remote Assistance in Virtual and Augmented Reality”, Proc. of the 28th Annual ACM Symposium on User Interface Software & Technology, pp.405-415 (2015).
[7] T. Piumsomboon, A. Day, B. Ens, Y. Lee, G. Lee and M. Billinghurst, “Exploring enhancements for remote mixed reality collaboration”, SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, pp.1-5 (2017).
[8] StateOfVirtualEvents2021, available from <https://s3.amazonaws.com/media.mediapost.com/uploads/StateOfVirtualEvents2021.pdf> (accessed 2024-1-19).
[9] C. Misa, D. Guse, O. Hohlfeld, R. Durairajan, A. Sperotto, A. Dainotti and R. Rejaie, “Lessons Learned Organizing the PAM 2020 Virtual Conference”, ACM SIGCOMM Computer Communication Review, Vol.50, Issue 3, pp.46-54 (2020).
[10] T. J. J. Tang and W. H. Li, “An Assistive EyeWear Prototype that interactively converts 3D Object Locations into Spatial Audio”, Proc. of ISWC’14, pp.119-126 (2014).

(Received: October 29, 2023)
(Accepted: January 29, 2024)

Yoshia Saito received his Ph.D. degree from Shizuoka University, Japan, in 2006. He had been an expert researcher at the National Institute of Information and Communications Technology (NICT) from 2004 to 2007, in Yokosuka, Japan. He was a lecturer from 2007 to 2011 at Iwate Prefectural University and he is currently an associate professor at the University. His research interests include computer networks and internet broadcasting.

Kei Sato received his master's degree from Iwate Prefectural University, Japan, in 2022. His research interests include 360-degree internet live broadcasting.