Regular Paper

Proposal of a Spherical POV Heatmap using Mixed Reality in 360-degree Internet Live Broadcasting

Yoshia Saito* and Junichiro Suto*

*Graduate School of Software and Information Science, Iwate Prefectural University, Japan

y-saito@iwate-pu.ac.jp

Abstract 360-degree internet live broadcasting enables viewers to change their point of view (POV) while watching the 360-degree live video and has an issue that the broadcaster cannot check the viewers' POV. To solve the issue, we have proposed a spherical POV heatmap using Augmented Reality (AR) on a smartphone so that the broadcaster can be aware of the viewers' POV. Although the spherical POV heatmap reduced communication errors, the response time increased due to the need to check the heatmap on a smartphone. In this paper, we propose a new spherical POV heatmap using Mixed Reality (MR) through an MR headset. The proposed system displays the heatmap in real space through the MR headset and reduces the response time by eliminating the need to check the smartphone. We implemented a prototype system and found it could reduce the response time through the evaluation.

Keywords: 360-degree Internet Live Broadcasting, Viewers' POV, Mixed Reality.

1 INTRODUCTION

360-degree video frees viewers from the constraints of viewing direction. The viewers can change their point of view (POV) and enjoy all directions of the video. Recently, 360-degree video has also been introduced in internet live broadcasting, where the viewers can typically communicate with the broadcaster via text chat. However, there is an issue that the broadcaster cannot check the viewers' POV in the 360-degree internet live broadcasting. Since the broadcaster does not know what the viewers saw and commented on, it can cause communication errors between the broadcaster and the viewers.

To solve the issue, we have proposed a spherical POV heatmap using Augmented Reality (AR) technology [1][2] (hereinafter called AR spherical POV heatmap). The AR spherical POV heatmap displays a spherical heatmap of the viewers' POV on a QR code using AR technology through a smartphone screen. The broadcaster checks the AR spherical POV heatmap and can know which direction the viewers are watching. Although it can reduce communication errors between the broadcaster and the viewers, it increases the response time to viewer comments due to the need to check the heatmap on the smartphone. The response time should be short to realize smooth communication between the broadcaster and viewers.

To overcome the disadvantage of the AR spherical POV heatmap, we propose a new spherical POV heatmap using Mixed Reality (MR) technology (hereinafter called MR spherical POV heatmap). The MR spherical POV heatmap displays a spherical heatmap of the viewers' POV on top of the omnidirectional camera in real space using the MR technology through an MR headset. Since the broadcaster can check the MR spherical POV heatmap in real space, it is expected that the response time to viewer comments can be shortened.

The contributions of this paper are summarized as follows:

- We developed and evaluated a prototype system of the MR spherical POV heatmap using HoloLens 2.
- We clarified that the MR spherical POV heatmap enabled the broadcaster to reduce the response time to viewer comments.

The rest of this paper is organized as follows. Section 2 describes related work and the AR spherical POV heatmap in the 360-degree internet live broadcasting as our previous work. Section 3 describes an overview of our proposed system. Section 4 describes the implementation of the prototype system. Section 5 describes an evaluation experiment to clarify the effects of the proposed system. Section 6 summarizes this study.

2 RELATED WORK

In this section, firstly we describe the role of eye gaze in human communication and the need to grasp viewers' POV for the broadcaster. Then, we describe the effectiveness of introducing MR technology in remote communication. At last, we show the detail of the AR spherical POV heatmap in our previous study and its issues.

2.1 Role of Eye Gaze in Communication

Many studies have described the importance of the eye gaze in human communication. Roel [3] proposed the GAZE Groupware System which was a study on the eye gaze information in human communication. This research verified the transmission of non-verbal information in a multi-participant teleconference system. He studied whether natural communication can be performed by conducting a meeting with non-verbal information in a virtual conference room. He discovered an issue that it was difficult to present eye gaze information in the system. This study concluded that who talked about what with whom would be analyzable if it was possible to show the eye gaze directions of the communicatees. David [4] also found the eye gaze information was an important factor that affected task performance in cooperative work. From these studies, the eye gaze is important information to express the communicatee's intentions and essential information in remote communication. ClearBoard [5] is a shared drawing medium which realizes a seamless shared drawing space and eye contact to support real-time remote collaboration by two users. They found "gaze awareness" is a most important feature for the collaboration.

The eye gaze information has been often represented on the flat display two-dimensionally. Angelo [6] investigated how remote pairs made use of gaze cursors which tracked their eye gaze on the display during a tightly coupled collaborative task. They found that the remote pairs used the gaze cursor to circumscribe the referential domain and they were also able to coordinate by indicating with both gaze and language to ground on the pieces. Xu [7] designed a hybrid meeting system which used an omnidirectional camera in the meeting room so that remote participants could look around. It also showed local participants where the remote participants' gazes were directed on a display such as a tablet PC under the omnidirectional camera. They found that it could provide powerful social cues, in ways similar to that of real gaze.

Several studies represent the eye gaze information in the real space three-dimensionally. OmniGaze [8] is a method for three-dimensionally displaying gaze information in telepresence. In this method, an omnidirectional camera is covered by a LED matrix display, and the lighting of the LED indicates the gaze information of a remote user. From the results of the evaluation experiment, it was clarified that the light information of the LED display on the sphere surface was effective for presenting the gaze information of a remote user. ThirdEye [9] is an add-on eye display that shows a remote participant's gaze direction. It improves the gaze estimation accuracy compared to the case where the remote participant's face is shown on a flat display. It indicates eye gaze information as 3D information in the real space is more effective than 2D information on a flat display.

Since the viewers' POV indicates directions where the viewers are watching centered on an omnidirectional camera, it has similar roles to eye gaze in remote communication. In this study, the viewers' POV is used for the same meaning as the eye gaze. Moreover, since eye gaze information should be shown as 3D information in the real space for remote communication, we try to introduce MR technology to show the information.

2.2 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 studies show the effectiveness of MR in remote communication between users.

Lee [10] 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 [11] studied the effect of MR guidance. An experiment was conducted to understand how to provide explicit spatial information in a collaborative MR environment. From the experiment, the result showed the MR guidance realized effective referencing through deixis.

Several studies also show the effectiveness of the hologram for remote communication [12-14]. From the related work, the improvement of communication between the broadcaster and viewers can be expected by introducing the MR technology to the POV heatmap in 360-degree Internet live broadcasting.

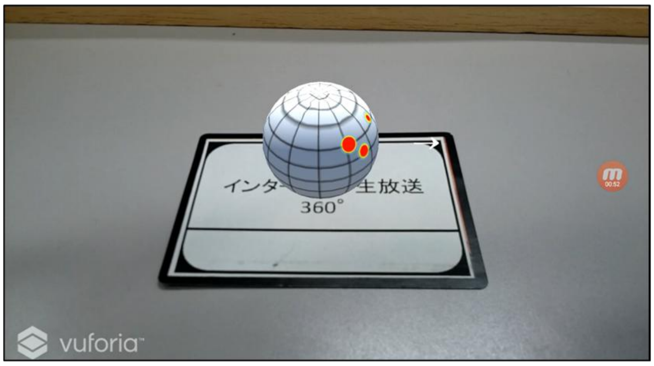

Figure 1: AR spherical POV heatmap through smartphone display

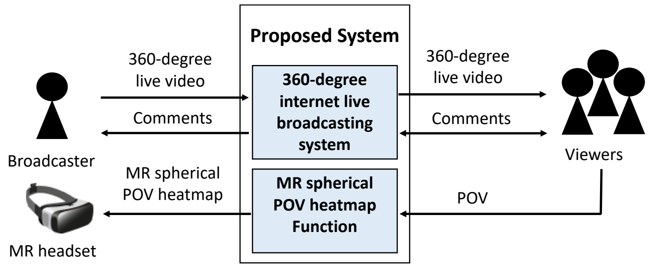

Figure 2: System model of the proposed system

2.3 AR Spherical POV Heatmap

The AR spherical POV heatmap shows a sphere that represents the broadcasting space on an AR marker through a smartphone as shown in Fig. 1. The spherical heatmap visualizes the viewers' POV by displaying the angular coordinate vector on the spherical surface as a heatmap. By synchronizing the direction of the image taken by the omnidirectional camera with the vector of the spherical heatmap, the broadcaster can grasp the viewers' POV in real space simply by checking the heatmap. In addition, the heatmap is easy to visualize multiple data at the same time. It is also possible to check the POV of multiple viewers. By checking the density pattern of the viewers' POV on the

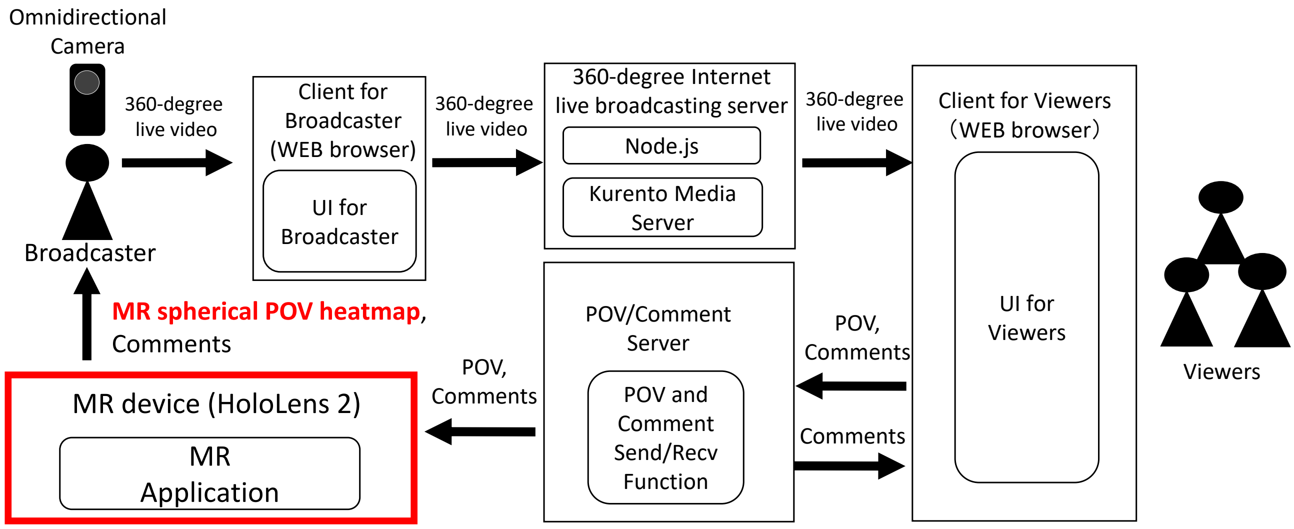

Figure 3: System architecture of the prototype system

heatmap, the broadcaster can grasp where many of the viewers are interested, and where the other viewers of the minority are watching. Therefore, it gives some hints about the viewers' interests and achieves smooth communication between the broadcaster and viewers.

In the evaluation experiments, the AR spherical POV heatmap could reduce communication errors by visualizing the viewers' POV and the speaking timing is easier for the broadcaster to grasp when waiting for the viewers' attention. On the other hand, the time it takes for the broadcaster to respond to the viewer's comment was increased by approximately 10 seconds compared with the broadcasting without the AR spherical POV heatmap. This result was due to the time required to check the heatmap through a smartphone. If the wasted time can be eliminated, the effectiveness of the spherical POV heatmap can be improved.

3 PROPOSED METHOD

In our previous study [1][2], the issue was that checking the viewers' POV was a burden for the broadcaster. This is because the spherical POV heatmap was implemented using AR and forced broadcasters to check their smartphone to see the heatmap. To solve this issue, we propose a system that displays the spherical POV heatmap using MR technology through an MR headset. The proposed system is expected to eliminate the need to confirm the POV heatmap using a smartphone in the previous study. The use of a head-mounted type MR device also solves the issue of both hands being occupied and reduces the burden on the broadcaster in 360-degree internet live broadcasting.

Figure 2 shows a system model of the proposed system. The proposed system consists of a 360-degree internet live broadcasting system and an MR spherical POV heatmap function. The broadcaster sends a 360-degree live video to the proposed system and the viewers watch the live video. The viewers send comments to the proposed system, which are sent to the broadcaster and the viewers. The proposed system receives viewers' POV in real-time and makes an MR spherical POV heatmap. The MR spherical POV heatmap is provided to an MR headset of the broadcaster.

The proposed system is expected to have several advantages by reducing the burden on the broadcaster as follows: (1) the time it takes for the broadcaster to respond to the viewer's comment is expected to be shorter than the previous study with few communication errors, (2) the speaking timing is easier for the broadcaster to grasp than the previous study when waiting for the viewers' attention. The advantages of the previous study will be further enhanced, and the disadvantage is solved by the proposed system.

4 IMPLEMENTATION

We implemented a prototype system of the MR spherical POV heatmap using HoloLens 2. In this section, we describe the architecture of the prototype system and its main application.

4.1 System Architecture

The prototype system was implemented by replacing the AR spherical POV heatmap in the previous study with an MR application for the MR spherical POV heatmap. Figure 3 shows the system architecture of the prototype system. The red square in the figure shows new implementation in this study and the other parts are diverted from the previous study. In the prototype system, we used Microsoft HoloLens 2 as an MR headset for the broadcaster and THETA V as an omnidirectional camera.

A broadcaster can start 360-degree internet live broadcasting using the broadcaster client on a web browser. The 360-degree internet live broadcasting server distributes it to viewers. The viewer can watch the 360-degree live video using the viewer client on a web browser and send text comments to the comment server. The comment server forwards the received comments to the viewers and the

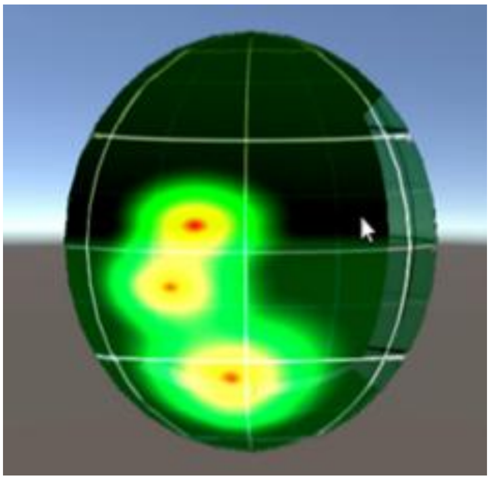

Figure 4: The heatmap when multiple viewers are watching from different POV

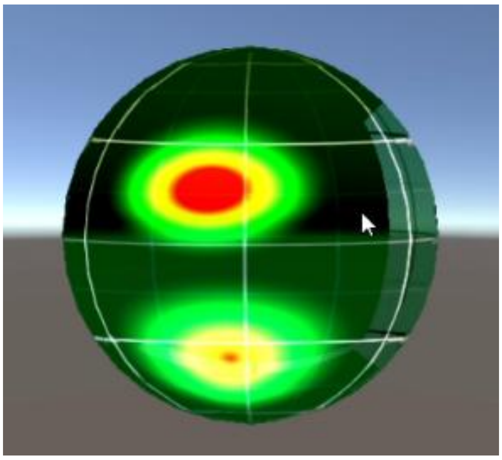

Figure 5: The heatmap when several viewers are watching in a particular POV

Figure 6: The user interface for the broadcaster through the HoloLens 2

broadcaster. The viewers can read the comments on their web browsers and the broadcaster can read the comments on the HoloLens 2. The viewer client also automatically sends the viewer's POV to the POV server every second. The POV is represented by spherical coordinates (r, θ, φ) where r denotes the radial distance, θ denotes the polar angle, and φ denotes the azimuthal angle. The POV server forwards the received viewers' POV to the MR application on the HoloLens 2.

4.2 MR Spherical POV Heatmap

The MR application on the HoloLens 2 is implemented by Unity and it shows a spherical POV heatmap and comments to the broadcaster. The MR application receives the viewers' POV and displays their heatmap on a sphere object which is shown above the omnidirectional camera. Based on a list of the viewers' POV received within a certain period, it creates circles for the heatmap. In this circle, the closer to the center, the higher the heat value, and the farther from the center, the lower the heat value. If a circle overlaps another circle, the heat values of the two circles are added and the overlapping area has a higher heat value.

Figure 4 shows an example of the heatmap when multiple viewers are watching the 360-degree live video from different POV. The broadcaster can see that the viewers' interests are dispersed in different directions. Figure 5 shows an example of the heatmap when several viewers' POV are concentrated. The broadcaster can see that several viewers are interested in a particular direction of the size of the red area.

Figure 6 shows the user interface for the broadcaster through the HoloLens 2. The MR application gets a coordinate and direction of the omnidirectional camera, and it displays the MR spherical POV heatmap which is fixed above the omnidirectional camera. The comment window tracks the broadcaster's sight so that the broadcaster can read the viewers' comments. The MR spherical POV heatmap is updated every second. The broadcasters can walk freely around the room wearing hololens2.

The backside heatmap color of the MR spherical heatmap is not visible unless the broadcaster goes around to the backside. This implementation was chosen because the information on the front heatmap overlaps with the information on the back heatmap, making it difficult to see.

In this study, we implemented this method to verify whether the display method used in the previous study, the AR spherical heatmap, could be improved using MR technology to reduce the effort required for confirmation. On the other hand, another implementation method that could be considered is to display the heatmap directly on the object itself in the direction the viewer is looking. However, issues such as the heatmap being difficult to see when it is outside the broadcaster's field of view were anticipated, so we decided to implement this method in the future.

5 EVALUATION

We conducted an evaluation experiment using the prototype system compared with the AR spherical POV heatmap. In this section, we describe the evaluation environment and the results.

5.1 Environment

The experiment was conducted under the same conditions as in the previous study [2]. The purpose of the experiment was to confirm whether the MR spherical POV heatmap reduced the burden on the broadcaster compared to the AR spherical POV heatmap.

Each experiment was conducted with one broadcaster and three viewers, for a total of four times. The role of broadcaster was performed by one different participant in the experiment at each time. This number of experimental participants was set because the use case in the previous study [2] assumes a single broadcaster and a small-scale internet live broadcast with less than 10 participants. The broadcasters and viewers are students at Iwate Prefectural University who have experience using computers but little experience using MR applications. The broadcaster and viewers were in different rooms in the experiment. The equipment used in the experiment was one notebook PC for the broadcaster, three laptop PCs for the viewers, one PC for the server, Ricoh Theta V for the omnidirectional camera, and HoloLens 2 for the MR headset.

The content of the broadcast was a chat about objects in the broadcaster's room. In the room, there were various objects. The viewers commented on objects in the broadcaster's room freely changing their POV. The procedure of the experiment is as follows:

- [Practice] The broadcaster and viewer practice the operation of the prototype system for 5 minutes.
- [Task 1] One of the viewers sends a designated question comment to the broadcaster 4 times at 3-minute intervals.
- [Task 2] The broadcaster directs the viewer's attention to a specified object and chats about it 3 times at 2-minute intervals.
- [Questionnaire] The viewers and broadcaster are given a 5-point scale questionnaire for each task.

In Task 1, the viewer communicates to the broadcaster by sending a question comment about an object in the room, such as "What kind of animal is this stuffed animal?". The text to be sent, the timing of the comment, and the object to which the viewer is directed are specified in the procedure manual. The broadcaster determines which object the viewer is commenting on and responds to the viewer's comment. If the broadcaster understands what the viewer is commenting on, the time it takes for the broadcaster to respond correctly to the viewer's comment is expected to be shorter. We measure the time and define it as the *response time*. We also count the *number of communication errors* if the broadcaster makes a mistake with the object the viewers are talking about.

In Task 2, the broadcaster points to a specified object and instructs the viewers "Please look at this". The broadcaster starts chatting with the viewers about the object when their POV are gathered. If the broadcaster is aware of the viewers' POV, it is expected to be able to start chatting at the same time when the viewer's POV is gathered. We measure the time between the broadcaster's attention instruction to the viewers and when the broadcaster begins to speak. It is defined as the *wait time*.

The viewers and the broadcaster complete a questionnaire after each task. After Task 1, the viewers were asked a

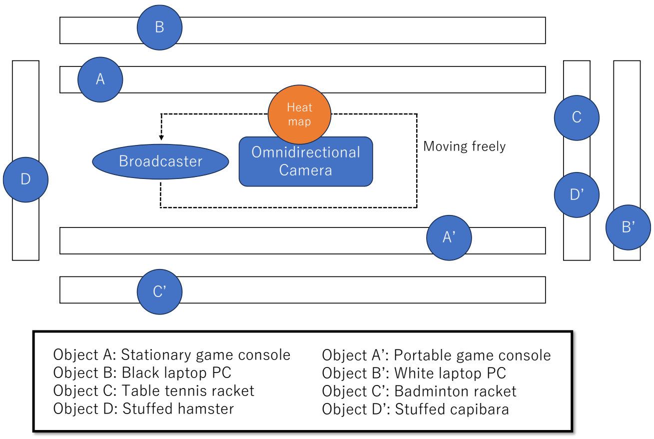

Figure 7: Location of all objects in the broadcaster's room for Task 1

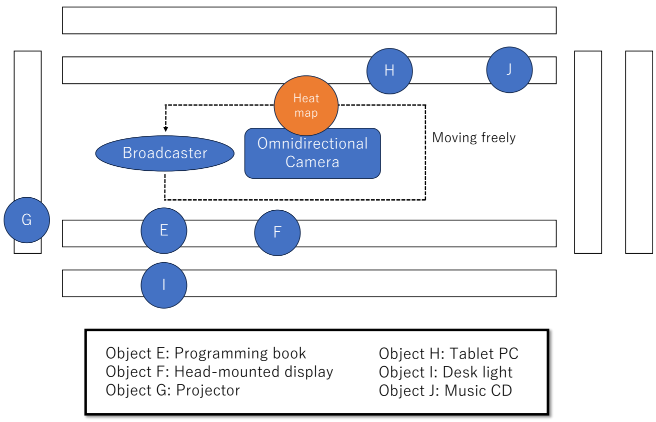

Figure 8: Location of all objects in the broadcaster's room for Task 2

questionnaire on a 5-point scale to determine how correctly they felt the comments were conveyed to the broadcaster. After Task 2, the broadcaster was asked a questionnaire on a 5-point scale to determine how well the broadcasters felt being able to capture the timing of when to begin speaking.

Figure 7 shows the location of all objects, the fixed position of the MR spherical heatmap and the initial position of the broadcaster in the broadcaster's room for Task 1. The MR spherical heatmap is displayed in a fixed position and does not move. The broadcaster can move freely within the area surrounded by desks and stand in front of the target object to speak. Four different types of objects (game consoles, laptop PCs, rackets, and stuffed animals) were prepared. Combinations using dashes, such as A and A', indicate target objects of the same type. For example, they are the same game console but differ in that one is stationary and the other is portable. This is to provide target objects of the same type that are difficult to distinguish when viewers refer to a particular game console. They are placed in different locations so that the same type of object was not in view at the same time. We also prepared other 6 objects (a programming book, a head-mounted display, a projector, a tablet PC, a desk light, and a music CD) in the broadcaster's room for Task 2 as shown in Fig. 8. The conditions such as the initial position of the broadcaster and the fixed position of the MR spherical heatmap are the same as in Task 1.

In this experiment, only the MR spherical POV heatmap was performed, and the environment of the previous study [2] was reproduced and the target objects were used same ones. The evaluation results of the AR spherical POV heatmap are reused from the previous study [2]. Therefore, since the participants in the experiments using the MR spherical POV heatmap and the AR spherical POV heatmap are different, there is no effect of experimental order on learning. Although there is a possibility that differences in results may arise due to the skills of the participants in each experiment, we assumed that there would be no significant differences since the participants in each experiment had similar skills.

5.2 Results in Task 1

Table 1 shows the results of measuring the response time in the evaluation experiment for Task 1 using the MR spherical POV heatmap. Table 2 shows the result of measuring the response time under a similar environment in the previous study. The mean response time was 21.47 seconds for the MR spherical POV heatmap, compared to 30.88 seconds for the AR spherical POV heatmap. In comparison, the MR spherical POV heatmap reduced the response time by about 9 seconds. We conducted a Mann-Whitney U test for the results. There was a significant difference in the mean response time to find ($p = 0.001586 < 0.05$). One of the reasons for this result is that it was necessary to confirm the POV heat map by using an AR application with a smartphone when confirming the POV in the previous study. However, this procedure could be omitted by displaying the MR spherical POV heatmap in real space, and the response time could be reduced.

Table 3 and Table 4 show the results of measuring the number of communication errors using the MR and AR spherical POV heatmap respectively. In terms of the number of communication errors, there were 4 communication errors in the MR spherical POV heatmap, while there was no communication error in the AR spherical POV heatmap. This is because some broadcasters selected the target object without checking the MR spherical POV heatmap at the first time. The number of communication errors could be reduced to zero if checking the MR spherical POV heatmap.

Table 5 and Table 6 show the results of the questionnaire "How correctly you felt the comments were conveyed to the broadcaster?" on a 5-point scale when the broadcaster used the MR and AR spherical POV heatmap respectively. The mean score was 3.7 in case of the MR spherical POV heatmap, while it was 3.56 in case of the AR spherical POV heatmap. We conducted a Mann-Whitney U test for the results and there was no difference between the mean scores ($p = 0.70626 > 0.05$). From the results of this questionnaire, the MR spherical POV heatmap can shorten the response time while keeping the accuracy of the communication with the broadcaster from the viewers' subjective point.

Table 1: The response time using the MR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
Object B'	15.34 sec	55.77 sec	16.04 sec	56.18 sec	21.47 sec
Object A'	41.92 sec	7.13 sec	27.16 sec	8.03 sec	
Object D'	14.69 sec	10.61 sec	31.23 sec	7.96 sec	
Object C'	11.61 sec	11.79 sec	16.78 sec	11.25 sec	

Table 2: The response time using the AR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
Object B'	45 sec	39 sec	46 sec	38 sec	30.88 sec
Object A'	29 sec	26 sec	29 sec	16 sec	
Object D'	35 sec	32 sec	24 sec	32 sec	
Object C'	29 sec	33 sec	13 sec	28 sec	

Table 3: The number of communication errors using the MR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Total
Object B'	0	2	0	2	4
Object A'	0	0	0	0	
Object D'	0	0	0	0	
Object C'	0	0	0	0	

Table 4: The number of communication errors using the AR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Total
Object B'	0	0	0	0	0
Object A'	0	0	0	0	
Object D'	0	0	0	0	
Object C'	0	0	0	0	

Table 5: The questionnaire result in case of the MR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
Object B'	4	3	5	2	3.7
Object A'	3	4	4	4	
Object D'	5	4	4	1	
Object C'	5	3	5	3	

Table 6: The questionnaire result in case of the AR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
Object B'	1	5	4	5	3.56
Object A'	4	3	4	4	
Object D'	4	2	5	4	
Object C'	3	3	3	3	

5.3 Results in Task 2

Table 7 and Table 8 show the results of measuring the wait time between the broadcaster's attention instruction to the viewers and when the broadcaster begins to speak in the evaluation experiment for Task 2 using the MR and AR spherical POV heatmap respectively. The mean wait time was 14.78 seconds for the MR spherical POV heatmap, compared to 16.67 seconds for the AR spherical POV heatmap. In comparison, the MR spherical POV heatmap shortened the wait time by approximately 2 seconds. Meanwhile, we conducted a Mann-Whitney U test for the results and there was no difference between the mean wait time (p = 0.20402 > 0.05). The MR spherical POV heatmap was not able to reduce the wait time although we expected to be able to reduce the time by eliminating the need to check the smartphone as with the previous result of the response time. One possible reason for this result is that when several heatmap circles were gathered into one heatmap circle, it was difficult to determine how many viewers were in the heatmap. In Task 2, the broadcaster provides an attentional instruction to a target object, and the viewer turns toward the viewing direction of the object. At this time, the heatmap circles are gathered into one heatmap circle, and the red area in the middle of the heatmap circle becomes larger. In the prototype system, when two heatmap circles overlap, the red area is clearly larger than when only one heatmap circle overlaps. However, when the number of overlaps increases from two to three, the red area is not as large as that from one to two, and it is difficult to understand the difference unless the broadcaster is accustomed to the display method. Since there were three viewers in the experiment, it was difficult to confirm that the POV of all the viewers had gathered to the target object, and this placed a burden on the broadcaster. Therefore, we consider that the wait time could not be reduced in the prototype system.

Table 9 shows the results of the questionnaire for the broadcaster, “How well you felt to be able to capture the timing of when to begin speaking?” on a 5-point scale when the broadcaster used the MR and AR spherical POV heatmap respectively. The mean score for the MR spherical POV heatmap was 3.25, compared to 4.5 for the AR spherical POV heatmap. The difference between MR and AR spherical POV heatmap was -1.25. We conducted a Mann-Whitney U test for the results and there was no difference between the mean score (p = 0.08326 > 0.05) while there was a trend. A possible reason for the lower score of the MR spherical POV heatmap compared to the AR spherical POV heatmap is that it is difficult to know how many viewers are there when several heatmap circles are gathered into one, as can be seen from the wait time result.

Furthermore, the broadcaster was asked to freely describe what they found advantages and what they found disadvantages about the prototype system. As an advantage point, many participants said that "it is easy to visually understand where the viewer is looking". As a disadvantage point, many participants said that "when the viewers’ POV is focused on one place, the heatmap becomes one circle, and it is difficult to grasp how many viewers are there".

Table 7: The wait time using the MR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
Object H	32.16 sec	4.8 sec	11.25 sec	9.03 sec	14.78 sec
Object I	14.9 sec	13.6 sec	12.56 sec	9.96 sec	
Object G	34.5 sec	15.2 sec	7.33 sec	12.05 sec	

Table 8: The wait time using the AR spherical POV heatmap

Target Object	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
Object J	10 sec	18 sec	24 sec	22 sec	16.67 sec
Object G	23 sec	15 sec	20 sec	20 sec	
Object F	12 sec	11 sec	13 sec	12 sec	

Table 9: The questionnaire result about the timing of when to begin speaking

Heatmap Type	1st Experiment	2nd Experiment	3rd Experiment	4th Experiment	Mean
MR	4	2	3	4	3.25
AR	4	5	4	5	4.5

There was also an opinion that "the conversation is delayed by one step because we have to look at the POV heatmap before speaking". This is because it is necessary to check the MR spherical POV heatmap above the omnidirectional camera before finding the target object. To solve this problem, instead of displaying the heatmap on a sphere, a method may be effective in which the heatmap is displayed directly on the target object.

6 CONCLUSION

We proposed the MR spherical POV heatmap to reduce the burden on the broadcaster in 360-degree internet live broadcasting, which displayed a spherical POV heatmap on the MR space. We compared it with our previous study, the AR spherical POV heatmap to clarify its advantages and issues. The results of evaluation experiments showed that the MR spherical POV heatmap reduced the response time compared to the AR spherical POV heatmap, and the objects could be found quickly. In addition, the use of the MR headset eliminated the need to check the smartphone and reduced the burden on the broadcaster. On the other hand, the wait time could not be reduced. One of the reasons for this result was that it was difficult to grasp how many viewers were there on the MR spherical POV heatmap. In future work, we will improve the way the heatmap circles are displayed so that it is more intuitive for the broadcaster to know approximately how many viewers are on the heatmap. Moreover, we will implement a method in which the heatmap is displayed directly on the target object.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number JP20K11794.

REFERENCES

[1] M. Takada, D. Nishioka and Y. Saito: "Proposal of a Spherical Heat Map in 360-Degree Internet Live Broadcasting Using Viewers' POV", The 2nd International Workshop on Mobile Ubiquitous Systems, Infrastructures, Communications and Applications (MUSICAL2019), pp.596-600 (2019).

[2] M. Takada, D. Nishioka and Y. Saito: "Proposal of the Viewer's POV Heat Map in 360-degree Internet Live Broadcasting", IPSJ Journal, Vol.62, No.1, pp. 145–159 (2021) (in Japanese).

[3] R. Vertegaal: "The GAZE groupware system: mediating joint attention in multiparty communication and collaboration", CHI '99 Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp. 294-301 (1999).

[4] D. Grayson and A. Monk: "Are you looking at me? Eye contact and desktop video conferencing", ACM Transactions on Computer-Human Interaction (TOCHI), Volume 10, Issue 3, pp. 221-243 (2003).

[5] H. Ishii and M. Kobayashi: "ClearBoard: a seamless medium for shared drawing and conversation with eye contact", CHI '92, Proc. SIGCHI Conference on Human Factors in Computing Systems, pp.525–532 (1992).

[6] S. D'Angelo and D. Gergle: "Gazed and confused: Understanding and designing shared gaze for remote collaboration", CHI Conference on Human Factors in Computing Systems, pp.2492–2496 (2016).

[7] B.X. Cornell, J. Ellis and T. Erickson: "Attention from Afar: Simulating the Gazes of Remote Participants in Hybrid Meetings", DIS '17, Proc. 2017 Conference on Designing Interactive Systems, pp.101–113 (2017).

[8] K. Shiro, A. Okada, T. Miyaki and J. Rekimoto: "OmniGaze: A Display-covered Omnidirectional Camerafor Conveying Remote User's Presence", HAI '18 Proceedings of the 6th International Conference on Human-Agent Interaction, pp.176-183 (2018).

[9] M. Otsuki, T. Kawano, K. Maruyama, H. Kuzuoka and Y. Suzuki: "ThirdEye: Simple Add-on Display to Represent Remote Participant's Gaze Direction in Video Communication", Proc. 2017 CHI Conference on Human Factors in Computing Systems, pp.5307–5312 (2017).

[10] 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).

[11] 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 CHI '21, pp. 1-14 (2021).

[12] O. Oda, C. Elvezio, M. Sukan, S. Feiner and B. Tversky: "Virtual Replicas for Remote Assistance in Virtual and Augmented Reality", Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, pp. 405-415 (2015).

[13] 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).

[14] S. Kim, G. Lee, W. Huang, H. Kim, W. Woo and M. Billinghurst: "Evaluation the Combination of Visual Communication Cues for HMD-based Mixed Reality Remote Collaboration", Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Paper No. 173, pp. 1-13 (2019).

(Received: December 27, 2023)
(Accepted: July 14, 2025)

Yoshia Saito received his Ph.D. degree from Shizuoka University, Japan, in 2006. He had been an expert researcher of National Institute of Information and Communications Technology (NICT) from 2004 to 2007, 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.

Junichiro Suto received his bachelor's degree from Iwate Prefectural University, Japan, in 2022 and he is currently a master's student in the university. His research interests include 360-degree internet live broadcasting and mixed reality.