Regular Paper

A Study on Analysis of Viewers' POV and Presentation to Broadcasters in Mobile 360-Degree Internet Live Broadcasting

Masaya Takada[*], and Yoshia Saito[*]

[*]Graduate School of Software and Information Science, Iwate Prefectural University, Japan
g236q002@s.iwate-pu.ac.jp, y-saito@iwate-pu.ac.jp

Abstract -360-degree Internet live broadcasting is a live broadcast using an omnidirectional camera. With the advent of various inexpensive omnidirectional cameras, this service is now available to general users. In addition, with the development of the Internet infrastructure, users are able to conduct the 360-degree Internet live broadcasting outdoors. This service can now be used as a mobile service. The features of this service have the ability to provide a greater amount of information than those of conventional broadcasting and a greater degree of freedom in viewing direction (POV: Point Of View). On the other hand, the 360-degree Internet live broadcasting services do not have the ability for the broadcaster to know the viewers' POV. The role of gaze information in remote communication is very important, as it shows the focus of the conversation and the object of interest. In other words, if the broadcaster cannot be aware of the viewers' POV, it is not possible to respond appropriately to the viewer's comments. For this problem, we have analyzed the POV and created an algorithm to detect viewers' interests. The algorithm used characteristics about the viewer's viewing behavior. It could detect significant POV changes which represent viewers' interests with 89.76% accuracy. In this paper, we show an experimental result to evaluate the effect of presenting the algorithm outputs to the broadcaster.

Keywords: 360-degree Internet live broadcasting, Viewers' POV.

1 INTRODUCTION

Interest in Internet live broadcasting services is increasing year by year, and many users enjoy real-time communication through live broadcasting services. Since Internet live broadcasting allows the broadcaster and viewers to enjoy real-time communication, it is also used as a communication tool. YouTube and some Internet live broadcasting services, such as Facebook, support for omnidirectional cameras. This service is called 360-degree Internet live broadcasting, and it is a groundbreaking service that provides more information about around of the broadcaster than that of the conventional broadcasting. For example, the broadcaster who goes sightseeing can provide the entire space of tourist spot using the 360-degree internet live broadcasting. However, unlike conventional broadcasting, the 360-degree Internet live broadcasting uses an omnidirectional camera. It is difficult for the broadcaster to intuitively grasp the viewer's viewing range (POV: Point Of View). It is necessary to present the viewers' POV to the broadcaster for smooth communication. This is because previous researches about remote communication have shown that the communicator's gaze indicated the target of interest or center of the topic [1]-[2]. The lack of information may lead to discrepancies in communication. The current services only support comments, when the viewers communicate with the broadcaster. The broadcaster has to respond based on the text information. Therefore, we address this issue by using the viewers' POV to facilitate communication by adding the ability to analyze and present it to the broadcaster.

The contributions of this paper are summarized as follows:

- We developed an algorithm to detect viewers' interests based on the characteristics of viewers' POV.
- We clarified the effects of the proposed algorithm when the analyzed results of the viewers' POV are presented to the broadcaster.

The rest of this paper is organized as follows. Section 2 describes the use cases and the advantage of our proposal. Section 3 describes the problem definition and approach in this study. Section 4 describes the related work and discuss the necessity of this study. Section 5 describes our 360-degree Internet live broadcasting system and its implementation. Section 6 describes the hypotheses about the viewer's viewing behavior and their testing. Section 7 describes an overview of the algorithm and its preliminary experiment. Chapter 8 describes improvements of the algorithm. Section 9 describes the effects of presenting the results of the POV analysis to the broadcaster. And we describe the results of the additional implementation and evaluation experiments. Section 10 describes a discussion and Section 11 summarizes this study.

2 USECASE

The use case which we envision for this research is that a single broadcaster delivers the situation of walking through a tourist spot. The broadcaster will visit a tourist spot and report about the spot to the viewers. The viewers can also request a report to the broadcaster using comments and they will be able to post their impressions of the broadcasting. The equipment used for the broadcast shall be a laptop computer and an omnidirectional camera. The broadcaster must carry a backpack with a camera mounter. Figure 1 shows the broadcaster who carry a backpack. The 360-degree video is centered around the upper part of the broadcaster's back. In 360-degree Internet broadcasting, the broadcaster will have access to more different means of communication than that of conventional broadcasting. For example, he/she can ask to direct attention to the object, or ask viewers to look for something from their surroundings. The broadcaster's motivation to use it is to make the walk in the tourist spot better. The idea is that the broadcaster can gain a virtual companion from the viewers, even if the broadcaster is traveling alone. In addition, the viewers' motivation to use it is to experience the virtual tourism without time and space constraints. The 360-degree Internet live broadcast provides a highly immersive experience as it allows the viewers to watch the full-sky image in any direction. The viewers can get a real sense of the sights as if they were there. Furthermore, the viewers can enjoy the tourist spot without knowledge about that place by taking advantage of the broadcaster’s tour because the broadcaster may have some kind of objective for the tour and act like a tour guide. On the other hand, the problem of the broadcaster's inability to grasp the viewers' POV is synonymous with inability to grasp the companion’s gaze. It may prevent smooth communication between the broadcaster and the viewers.

3 PROBLEM DESCRIPTION

In the use case, we envision the use of live 360-degree Internet broadcasting for the broadcast environment. In the 360-degree Internet live broadcasting, the broadcaster is required to provide broadcasting contents equivalent to the conventional ones under the circumstances in which they are hardly able to grasp the state of the viewers. Since the viewers can freely change direction of the view angle, they can get various information from all the directions of 360-degree angles according to their own interests. Simultaneously, the broadcaster has to respond to their various interests. In non-360-degree Internet live broadcasting, the viewers’ POV corresponds to the direction of the camera lens and all viewers watch the video at the same angle. The broadcaster can understand what the viewers watch and control the angle of the video arbitrarily by change the direction of the camera. On the other hand, the direction of the omnidirectional camera lens does not show the viewers’ POV. This issue has negative effects on communication between the broadcaster and the viewers.

When viewers find something interesting in the broadcaster's surroundings, they make comments to the broadcaster that include pronouns and so on. In linguistics, this behavior is referred to as "deixis" and the response to an instructional expression is expected to be a "Anaphora" that indicates the object corresponding to the instruction [3]. If these confirmations are not properly performed, the correct understanding of the context of the comment from the viewers may be difficult to understand and communication errors may occur. Therefore, we study ways to reduce communication errors based on research on the role and use of non-verbal information in the field of communication.

Figure 1: Equipment of the broadcaster.

4 RELATED WORK

There have been many studies on the role of non-verbal information in communication. In particular, gaze information has been shown to play an important role in communicating mutual intentions. The GAZE Groupware is a study of gaze information in communication [1]. In this study, the non-verbal information of the remote communication in a teleconference system is analyzed. He verified whether natural communication can be performed by conducting a meeting with nonverbal information in a virtual conference room. In addition, he discovered a problem that it is difficult to present gaze information because the space in which the conference participants reside is different in the remote meeting systems. He concludes that it is possible to analyze who talks about what by talking about the gaze directions of the communicatees.

Another study on mutual gaze in remote communication using videoconferencing systems [2] has revealed some interesting findings. The authors argue that the eye contact information of the communicatee is an important factor in the outcome of collaborative work with remote communication. Furthermore, the study also examined the method of presenting gaze information and concluded that the presentation of images including the eyes of the communicatee requires a certain size of images. In 360-degree Internet live broadcasting, the POV is the information that indicates the viewing direction and viewing range of the viewers, and it plays the same role as the gaze in remote communication.

On the analysis of viewers' POV in 360-degree video, a study of Yen-Chen Lin et al. examined on the correction of viewing direction in 360-degree video [4]. In this study, they examined a method of correcting the viewer's direction to the direction of the main story of a 360-degree video. They have implemented and evaluated two patterns of corrections: an automatic correction function and a correction with annotations. The results showed that there were multiple purposes and patterns in the viewer's viewing behavior and emphasize the need to analyze the viewer's viewing direction to provide a higher quality viewing experience.

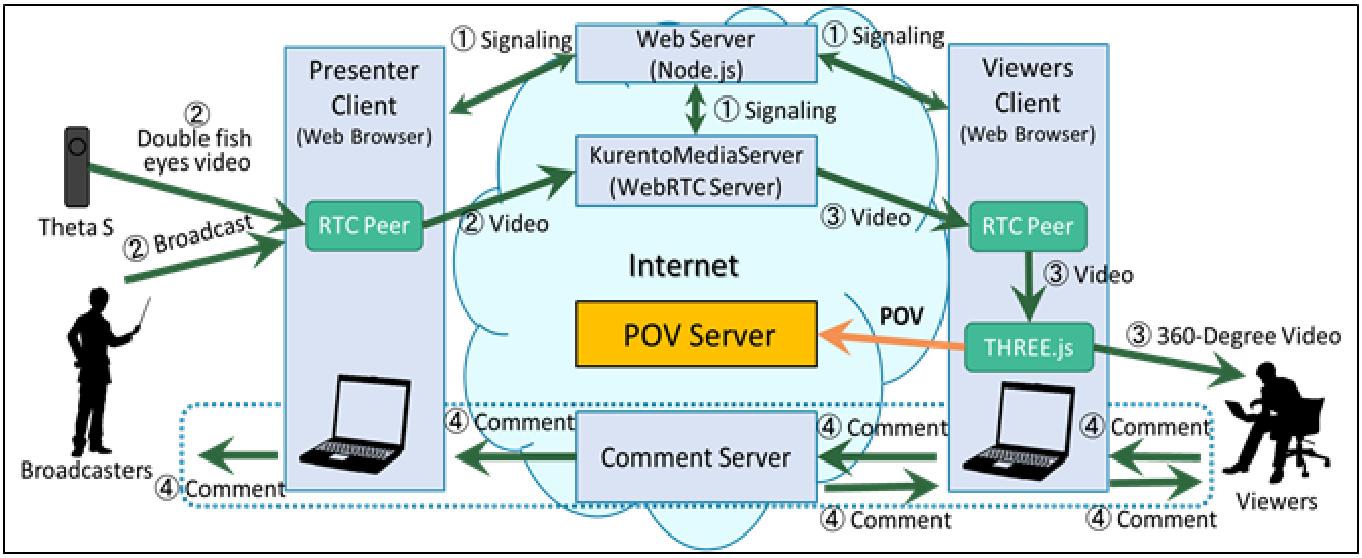

Figure 2: The system architecture of the 360-degree Internet live broadcasting system

YouTube provides a heat map analysis function for posted 360-degree videos, and the results of the analysis of the entire 360-degree videos are also available [5]. An analysis of the viewer's POV during viewing revealed the characteristics of watching a 360-degree video. The viewer's POV was directed most toward the 90-degree horizontal range centered on the front of the video, where 75% of the playback time was spent. It was also shown that only 20% of the users watched the full 360-degree range, even for the most popular videos.

5 INTERNET LIVEBROAD CASTING SYSTEM

The experiment environment consists of two components which are a 360-degree Internet live broadcast system and a POV server which collects the viewers' POV information. Figure 2 shows the architecture of the system. We use the Ricoh Theta S as an omnidirectional camera for the system.

The broadcaster accesses a broadcasting system and starts broadcasting by a web browser. Viewers access the broadcasting system and watch the live broadcasting. The viewers' POV information is sent to the POV server periodically.

Various well-known online video distribution services such as YouTube shift from Adobe Flash to HTML5 [6] for the video streaming. We also implemented a 360-degree Internet live broadcast system by adopting WebRTC that realized the video streaming on HTML5. We used the Kurento Media Server [7] for the video streaming server. The WebRTC is an API to provide real-time communication functions such as voice communication and video distribution without requiring any plug-in and installation of special software. We used a javascript library called Three.js [8] for processing images. The image acquired from the omnidirectional camera is mapped to a spherical object by using the library. The viewers' POV are managed by the angular coordinates of the two axes which are acquired every 100ms and the data is sent to the POV server. In this system, the POV is represented by the polar coordinate (Φ,θ,r) of the spherical surface. The POV server collects the POV information received from the viewers and saves the collected data in the database. Web server and POV server were implemented using node.js [9].

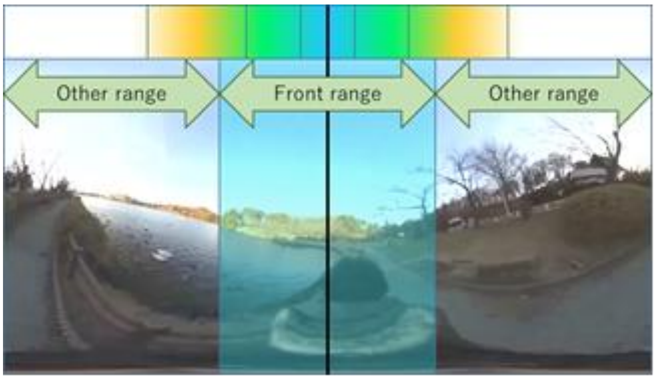

Figure 3: The equirectangular video and the front range.

6 HYPOTHESIS TESTING

6.1 Hypothesis

Based on the results of the related work, we have built three hypotheses concerning the characteristics about the viewers' POV in a 360-degree Internet live broadcasting. The first hypothesis is that "The viewers' POV is concentrated on the direction of the broadcaster's way in mobile environment". In the case of on-demand 360-degree videos, the object matter is displayed in the frontal direction because the video contributor will take or edit the video so that the viewers can fully enjoy the object matter of the video without change of POV. However, the 360-degree Internet live broadcasting is in real time and cannot be edited. Therefore, it is not possible to set an explicit frontal direction to the video, and the viewers may understand that the direction of the broadcaster's way can be the frontal direction. The second hypothesis is that "If the viewers' POV directs at other direction except the direction of the broadcaster's way, the viewing behaviors have meanings and there are some interesting objects for the viewers in the direction". In a 360-degree Internet live broadcasting, viewers can change the

POV according to their own interests. As shown in related work [5], viewers are most likely to watch in the frontal direction, and when they watch in directions other than the frontal direction, their interests and concerns are likely to be directed in that direction. The third hypothesis is that "The viewers' POV returns to the direction of the broadcaster's way after the viewers' interests are satisfied". If the viewer's interest is satisfied or the target of interest is no longer visible to the other direction, they no longer need to view to the other direction and return their viewing to the frontal direction. Summarizing the above hypothesis, in a 360-degree Internet live broadcasting, the viewer's POV is directed in the direction where the broadcaster is going. It changes from the frontal direction to the other direction when a target of interest is found. Thereafter, when the interest is satisfied or the target is no longer visible, the viewer's POV is expected to return to the direction where the broadcaster is going.

6.2 Verification Experiment

We conducted a broadcasting experiment to test these hypotheses. The purpose of the experiment is to test the three hypothesis and collect the data needed to create an analysis algorithm. Six collaborators participated in the experiment, and three collaborators each played the role of a viewer, and the experiment was conducted twice. The role of the broadcaster was played by a same member of the research team. The location of the broadcast is Takamatsu Pond in Morioka City, Iwate Prefecture, which is famous as a place where swans fly and a famous place for cherry blossoms. We made a 30-minute broadcast while moving around the pond, once for right and once for left. We explain to the collaborators that the purpose is to test the operation of a 360-degree Internet live broadcast. The POV was stored in the POV server every 100 milliseconds, and the hypothesis was tested by analyzing the POV log after the broadcast.

Since the POV logs could store 18,000 data per person in a 30-minute broadcast, 1,080,000 data were collected in two broadcasts. The analysis showed that the time when the POV was directed to a 90-degree range centered on the direction of the broadcaster's path was 75.89% in the first experiment and 80.33% in the second experiment. It was found that the viewer's POV was longest directed at a 90-degree range centered on the broadcaster's direction of walk, as explained in Hypothesis 1. Since the results of the analysis were close to those of the YouTube report, and we decided to proceed with the data analysis based on the collected data. From this point, we followed the YouTube report [5] and called a 90-degree area centered on the broadcaster's direction of walk as the front range. And the ranges other than the front range will be called the other range. Figure 3 shows the equirectangular video and the front range. In addition, we confirmed that Hypothesis 2 was valid because we found several cases in which the target of the viewer's interest was the same as the POV when the POV leave from the front range. We confirmed that the return of the POV to the front range an average of 11.41 seconds after the POV leaves from the front range. This confirms that Hypothesis 3 is also valid.

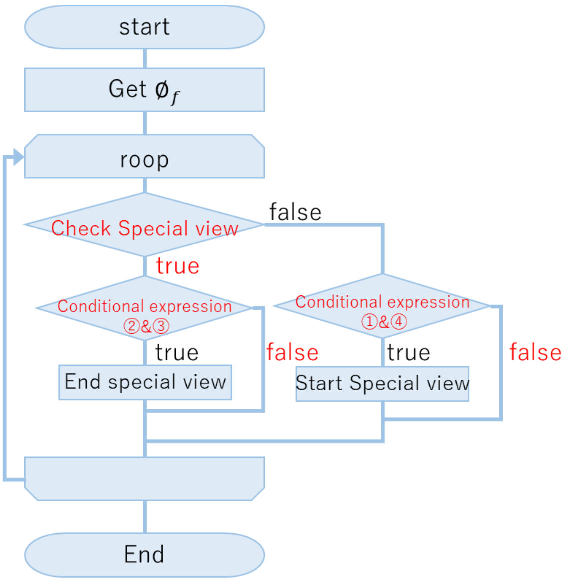

Figure 4: Flowchart of the algorithm

The variable
$\emptyset_i = \{\emptyset_{i1}, \emptyset_{i2} \dots \emptyset_{in}\}$
Horizontal polar coordinates of viewer POV
forward = $\emptyset_f$
Horizontal polar coordinates broadcaster's way

Conditional expression①
forward - 45 < $\emptyset_{i\,n-9}$ < forward + 45
Conditional expression②
$\emptyset_{i\,n-9}$ < forward - 45 || forward + 45 < $\emptyset_{i\,n-9}$
Conditional expression③
forward - 45 < $\emptyset_{i\,n}$ < forward + 45
Conditional expression④
$\emptyset_{i\,n}$ < forward - 45 || forward + 45 < $\emptyset_{i\,n}$

Figure 5: The variables and conditional of flowchart.

6.3 Algorithm Examination

The algorithm acquires the direction of the broadcaster's way as horizontal coordinates from 0 to 360. Then, based on the latest POV data for 10 cases, the algorithm decides which of the four states corresponds to which one of them is applicable, using a conditional expression. However, since the POV data is acquired every 100 milliseconds, immediate state transitions would lead to many false positives. For this reason, we adopt the stacking method for state transitions and do not transition until 10 different states are input. Also, the number of inputs is reset for each of the states that have been entered when the state transition occurs. The algorithm is able to deter the false detection of users viewing near the front range boundary. On the other hand, algorithm is not possible detection for other range viewing that less than 1 second, because it becomes possible to detect it 1 second after start of other range viewing. However, data analysis confirms that other range viewing takes place for an average of 11.41 seconds It is believed that it is possible to detect a sufficient. Due to the nature of the omnidirectional camera, it is not clear how the camera will be installed and fixed, so it is expected that the direction of the broadcaster's way will be different for each broadcast. In order to implement the algorithm, you need to acquire and compensate the direction of the broadcaster's way using acceleration sensors.

7 ALGORITHM

7.1 Overview

By testing three hypotheses, we found that viewers changed their POV according to their own interests when the POV was directed to the other range. Therefore, we developed an algorithm to detect POV viewing within other ranges. We determined the classification of the viewer's state and the buffer size to be used in the detection algorithm based on the data used to test the hypothesis. The viewer's state is classified into the following four categories. The state in which the viewer is viewing the front range is called the "normal viewing". The state in which the POV changes from the front range to other range is called "start of other range viewing". The state in which the viewer is continuously viewing the other range is called "other range viewing". The state that returns to the front range is "end of other range viewing". The buffer size for the analysis was set to 10 data of POV. We compared the detection accuracy and immediate response, it was determined that this buffer size was the most appropriate for. Normal viewing is the state which the viewer's POV changes only within the front range, and we are expected to remain in this state for the longest period of time during the broadcast. The algorithm detects and analyzes the POV in the state of other range viewing by triggering the start of other range viewing and end of other range viewing. Figure 4 shows a flowchart of the algorithm we created. Figure 5 shows the variables and conditional expressions used in the flowchart.

7.2 Verification

We verified the algorithm using the collected data if it correctly detected the POV of other range viewing. As a result, algorithm detected the other range viewing 73 times of at the first broadcast, and detected 95 times at the second broadcast, for a total of 168 times. We compared each results of detection with the recorded video to see what was being viewed. There were 149 cases (88.69%) in which the target object was obvious, and there were 19 cases (11.30%) in which the target object was unclear. Detection results that were unclear on the target object were mainly operation checks and a search of the area. This confirms that the POV pointed the other range was manipulated to view something. We divided 149 data into two groups in terms of whether they can be used for communication, such as whether the objects can be used as topics for broadcasting. There were 76 (51.00%) cases that were judged to be useful for communication by the broadcaster and 73 (48.99%) cases that were judged to be difficult to use. There was no significant difference between the two classifications in terms of whether they can be used for communication.

Table 1: Number and target of broadcasting.

Broadcast	A Target: Yes Topicality: Yes	B Target: Yes Topicality: No	C Target: No
1st broadcast	35 data	29 data	9 data
	A rowing boat A small shrine near the pond A characteristic building	Pedestrians A flock of birds Swans on the opposite of the pond	Looking around Test operations
2nd broadcast	41 data	44 data	10 data
	A wild duck landing on the pond A vinyl house A person shooting a photograph of the pond	Trees by the roadside Bicycles parking around the pond Sunset	Looking around Test operations
Percentage	45.24% (76/168)	43.45% (73/168)	11.31% (19/168)

Those that can be used for communication are called Group A. Those that are difficult to use for communication are called Group B. Those that objects could not be identified are called Group C. Table 1 shows the number of the three categories and their targets for each broadcast. Group A's viewing objects were a small shrine built on the bank of a pond and a duck landing on the water that had flown in. Group B's viewing objects were cars and flocks of birds parked around the perimeter of the pond. The data for Group C was a confirmation of operations and a search of the area immediately after the start of the broadcast. After analyzing the data for each group, we found that average time of other range viewing of A, B and C were 9.76, 12.09 and 15.36 seconds respectively. The standard deviations of A, B and C were 7.89, 23.37 and 23.27. However, each of B and C had one data item which duration of other range viewing was more than 100 seconds and they might contain extreme outliers. When these outliers were removed, we found that time of other range viewing of A, B and C were 9.76, 9.73 and 10.44 seconds, respectively. The standard deviations of A, B and C were 7.89, 12.13 and 10.54. From additional interviews, we found that the viewers interrupted and left the POV operation to enter comments.

7.3 Improvements

The algorithm detected the POV of the viewers directed to other ranges. The significant data detected by the algorithm were 149 of 168 (88.69%) Data. However, the number of data classified as useful for communication was only 51.00% (76/149). In Group B, there were many detections that had already noticed by the broadcaster because the viewers were viewing near the boundary of the range. Therefore, it is necessary to re-examine the boundaries of the front range setting in this study. The POV is information that indicates in which direction the center of the viewing image is pointed. Therefore, when the object of viewing is at the edge of the screen, the POV is directed to the other range but viewers may view front range. In the 360-degree Internet live broadcasting system, the camera angle of view in three.js is set at 35 degrees. To completely hide the front range, it is necessary to point the POV at a range of ±62.5

degrees or more. We decided to redefine the boundaries of the algorithm to ±60 degrees to analyze the data again.

As a result, new algorithm detected other range viewing 127 times, and we found that Group A and Group B were 114(89.76%). New algorithm detected Group C 13(10.23%). Among the 114 data items, data classified as A and B were 70(61.40%) and 44(38.59%). By improving the algorithm, we were able to prevent false positives and improve the percentage of group A by about 10%. By extending the front range, 41 viewings near the boundary of the front range were excluded from the detection of other range viewing. 35 data that were excluded from the detection were classified as Group B or Group C. We were able to exclude from the detection those that were close to the boundary line, which were less topical, such as people looking at the pond, people passing by, and trees growing beside the sidewalk. On the other hand, 6 data classified in group A were excluded from the detection. Excluded from the detection were the pond management office building, a swan boat covered with a blue sheet, a passerby with a camera, and a duck landing. The reason for these being out of detection is that the broadcaster was standing still and could watch without having to move the POV significantly. Therefore, POV was included in the extended front range. The algorithm may be difficult to detect other range viewing when the broadcaster is at a standing still.

8 EVALUATION

To present the results of the algorithm's analysis, we implemented an additional POV indication function. We superimpose different colored circles on the equirectangular video of the broadcaster's user interface in order to show the viewers' other range viewing. The colors are different for each user, and the center of the circle shows the POV. Figure 6 shows the superimposed display of the POV. The diameter of the circle matches the horizontal 25 degrees of the video. The viewers are viewing a 75-degree horizontal angle of view through the viewer's user interface, but it is impossible to view the entire 75-degree range of the video. The display used in this experiment is 21.5 inches with a resolution of 1920 x 1080. It is full HD quality, and the video displayed on the viewer client is 14.28 cm x 10.71 cm. It is said that the field of view where a human being can process information accurately without moving his or her eyes and head is 5 degrees or less in the horizontal direction. If the distance between the viewer's eyes and the display is 60cm, the range is 5.24 cm or less in the horizontal direction. The viewer can watch only about one-third of the video in the horizontal direction. Therefore, the indication of the POV is based on a range of 25 degrees in the horizontal direction, which is one-third of the camera angle of view. We have decided to present them. This criterion is used as a test for this evaluation.

8.1 Evaluation Experiment

An experiment was conducted to present the results of viewers' POV analysis to the broadcaster using an improved algorithm. The broadcast experiment was conducted twice. The broadcast route, broadcast time and equipment were the same as in the hypothesis test. In this experiment, there were eight experiment participants, because the role of the broadcaster was also played by the experiment participants. The broadcaster was informed of the broadcast procedure and that the viewers' POV would be displayed on a laptop computer. They were also instructed to stop at the edge of the sidewalk when checking display of the POV. No instructions or physical restrictions are placed on the content of the broadcast. We only told the viewers that we would be conducting a 360-degree Internet live broadcasting test. The expected effects of the experiment are the improvement of

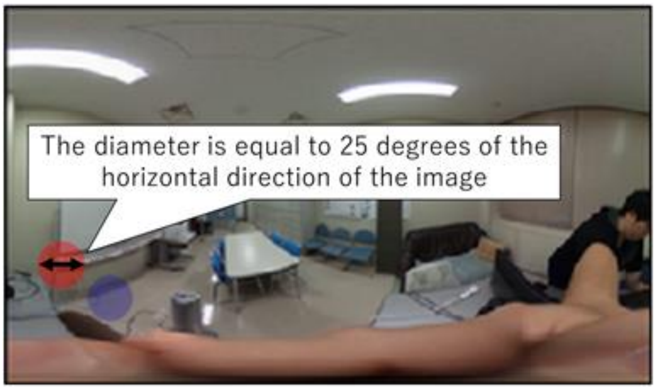

Figure 6: Superimposed POV indication

Table 2: The percentage results of the topical and notice.

broadcast	Topical	Notice	Persentage
1st broadcast	Yes：20 (57.14%)	14 (40.0%)	62.86% (22/35)
	No：15 (42.86%)	8 (22.86%)	
2nd broadcast	Yes：39 (84.78%)	22 (47.83%)	54.35% (25/46)
	No：7 (15.22%)	3 (6.52%)	

Table 3: The Results of target objects and reason

Broadcast	Topical	Target×Number	Reason
1st broadcast (35)	Yes (20)	Birds×9	They were Interesting and can expand the topic.
		Rowing boats×4	I turned out viewers' interests in boats.
		Edge of pond×3	The shape was easy to expand the topic.
		Other×4	I could expand the topic of dog walking.
	No (15)	Street tree×5	I thought it couldn't expand the topic
		State of pond×5	I didn't see anything in the pond.
		Buildings×4	It's a just landscape without features.
		Backward direction×1	There was nothing in particular behind me.
2nd broadcast (46)	Yes (39)	Frozen pond×11	The viewers were interested in it.
		Birds×9	They were Interesting and can expand the topic.
		Persons×9	A parent and a child enjoyed feeding birds and I thought it had topicality.
		Other×10	I could talked about a policeman who asked me questions, persons who took a walk and visited the pond.
	No (7)	Frozen pond×2	Birds were landing when talking about the birds.
		Cars×2	It's a just landscape without features.
		Building×1	It's a just landscape without features.
		Other×2	I felt it was a meaningless viewing.

understandability of the viewers' broadcasting needs and events which could not be noticed by the conventional broadcasting system. This is because the algorithm only extracts the necessary POV, which reduces the number of POVs that the broadcaster needs to check and makes it easier for the viewer's interest to be identified. After the experiment, we extracted data of other range viewing from the POV log and presented the broadcaster together with the recorded video for interview. For each of the other range of viewing, we asked whether they noticed the POV indication, whether they grasp the object from the POV indication, whether the object could be used as a topic, and the reason why it was topical. Additional questions were asked in an open-ended format if we had interested in the answers.

8.2 Evaluation Results

From the experimental results, the detecting other range viewing by the improved algorithm were Group A and Group B were 35 and 46. Table 2 shows the results of the interviews with the broadcasters after the experiment if there were the topicality and the reasons for the responses. We thought the broadcaster have to stop once to check the POV indication and might not notice the POV indication because he/she cannot watch the screen of the laptop during the walking. However, from the interviews, we found that the broadcasters frequently stopped to check their comments and were able to notice the POV indication. The number of times they noticed the POV indication was 62.86% (22/35) for the first broadcast and 54.35% (25/46) for the second broadcast. Among those noticed in the POV presentations, 14 were judged to be topical, and 22 were judged to be topical.

Table 3 shows the results of target objects of viewing and reason. In some cases of other range viewing, the viewers watched same objects at the same time. The ones shown in red are the representative targets of the viewing which were judged to be topical. In the case of presence of topicality, it was the pond or birds whose state was frequently changed by the flight of wild birds and the movement of carp. For example, it shows a main attraction of the pond such as a rowing swan boat covered with blue sheets. In the case of absence of topicality, objects whose state was not frequently changed, such as street trees and buildings, or objects which do not particularly catch your eye, such as passersby, were mentioned.

The broadcaster explained that these reason for this decision was that he was able to grasp the interest in wild birds and that he was able to get a lot of comments on the topic of police patrols where the POV was shown. Regarding the presentation of the POV which was not noticed during the broadcast, we also received comments that "if I had noticed the interest in the boats, I would have moved a little closer to the boats" and "if I had been able to notice it during the broadcast, I would have used it as a topic of discussion". Even for those which were judged not to be topical, the broadcasters commented that it did not require a big effort to check and did not bother them too much when presented, and that it was helpful to be able to notice the car traffic, even though it was not topical. Therefore, it is expected that the inclusion of non-topical other range of viewing in the POV indication will not have a significant negative impact on communication and broadcasting.

9 DISCUSSION

9.1 Effect of POV Presentation

The reasons for the presence or absence of topicality can be divided into three categories. The first reason is that the broadcaster feels topical that the state of the object or the change in condition of the object. The broadcaster can get a topic from the changing condition of the object. If the state of the object doesn't change, it is expected that it will be difficult to keep talking about it. In the experiment, the broadcaster talked about changeable objects such as wild birds and frozen ponds. The second reason is that it enables the broadcaster to confirm the interest of the viewers, which was not confirmed by the comments. The only means of communication from the viewer to the broadcaster is performed through text-based comments, and the broadcaster cannot understand interest of the viewers unless the viewers send comments about their interest on their own initiative. In the experiment, the broadcaster noticed that the viewers were interested in boats that were not mentioned in the comments, and could use them as topics of conversation. We also found that viewers may be interested in the same subject matter. The third reason is that the broadcaster can confirm viewers' interests about what the broadcaster said. In the experiment, although the broadcaster made frequent calls to viewers, the response comments was not that great. However, the broadcaster could confirm the viewers' interests of an object which the broadcaster talked about by checking the POV indication directed at the object of the topic.

9.2 Future Issues

From these results, by presenting the POV of other range viewing, the broadcaster can not only adapt the topic to the viewers' interests, but also get the response to the broadcaster's statement from the POV indication. Furthermore, three criteria potentially can be used to analyze the topicality of the POV. In half of the cases, we found that the broadcaster could not be aware of the displays of superimposing the POV on the broadcaster's user interface. It is necessary to consider a method of presenting the POV so that the broadcaster can check the POV indication even while walking. As a concrete method, we can use senses other than vision, such as sound and vibration. It is also necessary to consider a method for automatic identification of topicality because some POV are presented without topicality.

False positives from the analysis algorithm were classified and methods for improvement were discussed. The items which were detected incorrectly this time can be classified into three patterns. The first pattern was viewing near the boundary line and the POV left from the front range for a very short time. The countermeasure to the first pattern was to extend the front range. We think another countermeasure is to weight the data based on the distance from the boundary line. The second pattern is stopped the POV control and left in viewing of other range. The countermeasure to the second pattern is to filter by the amount of POV movement. If possible detection of interruptions in operation based on the viewer's POV change while viewing other ranges, it is possible to exclude them from the detection. The third pattern is an unexpected behavior such as checking operations. A possible countermeasure to the third pattern is to filter the POV operation direction. By converting the data, we can analyze the degree of meandering, and this would allow us to determine whether the object of viewing exists.

10 CONCLUSION

In this study, we investigated and improved a POV analysis algorithm that uses the viewers' viewing behavioral characteristics in a 360-degree Internet live broadcasting. Behavioral characteristics were validated and we implemented an algorithm that uses viewers' behavioral characteristics. We conducted initial evaluation of the algorithm was conducted, and we could detect the significant POV changes. Further improvement of the algorithm was investigated to increase the percentage of beneficial data which can be utilized for broadcasting, we found that significant POV with 89.76% accuracy. As a result of the experiment for the impact of presenting the analysis results, we found that POV indication made it easy for the broadcaster to understand interests of the viewers. In addition, we found that the inclusion of non-topical other range of viewing in the POV indication will not have a significant negative impact on communication and broadcasting.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number JP20K11794.

REFERENCES

[1] 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).

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

[3] S. C. Lee, "Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse", Massachusetts Institute of Technology201 Vassar Street, W59-200 Cambridge, MA, United States.

[4] Y.C. Lin, Y. J. Chang, H. N. Hu, H. T. Cheng, C. W. Huang, M. Sun, "Tell Me Where to Look: Investigating Ways for Assisting Focus in 360° Video", CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 2535-2545(2017).

[5] YouTube Creator Blog: Hot and Cold: Heatmaps in VR, https://youtube-creators.googleblog.com/2017/06/hot-and-cold-heatmaps-in-vr.html.

[6] YouTube Engineering and Developers Blog: YouTube now defaults to HTML5 <video>, https://youtube-eng.googleblog.com/2015/01/youtube-now-defaults-to-html5_27.html, November 7, 2018.

[7] Kurento, http://www.kurento.org/, November 7, 2018.

[8] three.js - Javascript 3D library, https://threejs.org/, November 7, 2018.

[9] Node.js, https://nodejs.org/ja/, November 7, 2018.

(Received October 29, 2020)
(Accepted December 2, 2020)

Masaya Takada. received his master degree from Iwate Prefectural University, Japan, in 2018. He is currently studying software engineering at the post-doctoral program at the University. His research interests include 360-degree Internet live broadcasting. He is a member of IPSJ and IEEE.

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. He is a member of IPSJ, IEEE, and ACM.