Proposal of a Spherical Heat Map in 360-Degree Internet Live Broadcasting Using Viewers' POV

Masaya Takada
Iwate Prefectural University Graduate School of Software and Information Science
Japan
g236q002@s.iwate-pu.ac.jp

Dai Nishioka
Iwate Prefectural University Graduate School of Software and Information Science
Japan
d-nishi@iwate-pu.ac.jp

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

***Abstract*—In this paper, we propose a viewers' POV (points of view) visualization method in 360-degree Internet live broadcasting using a spherical heat map. A critical problem of the 360-degree Internet live broadcasting is that the broadcaster cannot be aware of the viewers' POV which mean directions where the viewers are watching on the 360-degree video. In order to solve this problem, we use the spherical heat map which visualizes density of the viewers' POV on the sphere handling the POV data as vector data from the center point of the sphere. The spherical heat map enables the broadcaster to grasp viewers' POV easily and communicate with the viewers smoothly in the 360-degree Internet live broadcasting. This paper shows our prototype system of the spherical heat map for the viewers' POV and effectiveness of the prototype system by conducting an initial evaluation.**

Keywords—*360-degree Video, Internet Live Broadcasting, POV, Communication Support*

I. INTRODUCTION

In recent years, 360-degree video contents using omnidirectional cameras are getting popular. Various well-known web services such as YouTube, Twitter and Facebook support 360-degree Internet live broadcasting which broadcasts 360-degree video to viewers in real time via the Internet. There is an advantage of the 360-degree video which can provide more information to the viewers than the conventional videos. Therefore, it has high affinity with the contents to appeal the entire space such as tourist attractions. Since the Internet live broadcasting service is sometimes used for enjoying communication between the broadcaster and the viewers, it is serving the role as a communication tool. These services cannot provide sufficient non-verbal information although they can be used easily for the communication. The viewers and the broadcaster cannot understand each other in terms of their gaze directions in the 360-degree Internet live broadcasting. The viewers can check the state of the broadcaster. The broadcaster, however, has to guess the viewers' state only from their comments. Researches of the role of gaze in communication[1][2][3] indicate that the gaze information is necessary to understand communicatee's attention and the center of a current topic of conversation. It means the broadcaster cannot understand the viewers' interests without awareness of the viewers' POV (points of view) which mean directions where the viewers are watching on the 360-degree video. It prevents smooth communication between the broadcaster and the viewers in the 360-degree Internet live broadcasting. Therefore, it is necessary to allow the broadcaster to be aware of viewers' POV. In this paper, we propose a viewers' POV visualization method in 360-degree Internet live broadcasting using a spherical heat map. We show our prototype system of the spherical heat map for the viewers' POV and effectiveness of the prototype system by conducting an initial evaluation.

II. PROBLEM DEFINITION

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.

Many studies have been conducted on the important role of nonverbal information in communication. Particularly in the research on the role of gaze, it is clarified that the gaze of the communicatee is the information indicating the intention of remark. The GAZE Groupware System[1] is a research on the gaze information in communication. This research verifies the transmission of non-verbal information in a multi-participant teleconference system. 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.

978-1-5386-9151-9/19/$31.00 ©2019 IEEE

"OmniGaze"[4] is a method for three-dimensionally displaying gaze information of the presenter in telepresence. In this method, the omnidirectional camera device is covered by the LED matrix display, and the lighting of the LED indicates gaze information of the presenter. As a result of the evaluation experiment, it is clarified which the light information of the LED displayed on the sphere surface is effective for presenting gaze information of the presenter. From this research, we can see to present gaze information is effective for remote communication using 360-degree videos. Our research aims to present multiple viewers' POV (gaze information) to a single broadcaster and it is different from the OmniGaze in terms of the objectives.

In the 360-degree Internet live broadcasting, the broadcaster needs to communicate with the viewers and responds to their interests despite that the viewers' POV which offer a clue of the interests cannot be grasped. Such environment is inadequate for the communication and information of the viewers' POV must be offered to the broadcaster in order to achieve smooth communication between the broadcaster and viewers.

III. Description of Spherical Heat Map

Some researches create a heat map overlaid on an original 360-degree image of the equirectangular [5] to visualize viewers' POV. However, most of the image presented by the equirectangular image is not within the broadcaster's viewing angle. The broadcaster cannot understand direction of POV accurately unless he/she know the entire surrounding space in advance. We considered the spherical heat map is effective as a method to present viewers' POV three-dimensionally in real space.

The spherical heat map shows a sphere which represents the broadcasting space centering on an omnidirectional camera. 360-degree video contents are generally implemented by mapping an image of the equirectangular format or the dual fish eyes format to a sphere. The viewers can watch a part of the 360-degree image from various directions by rotating the camera placed at the center of the sphere in accordance with drag or swipe operations. The viewers' POV are managed by the angular coordinate vector of two variables indicated by φ and θ. The spherical heat map visualizes the viewer's POV by displaying the angular coordinate vector on the spherical surface as a heat map.

By synchronizing the direction of the image taken by the omnidirectional camera with the vector of the spherical heat map, the broadcaster can grasp the viewer's POV in the real space simply by checking the heat map. In addition, the heat map 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 heat map, 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 of the viewers' interests and achieves smooth communication between the broadcaster and viewers.

However, it is difficult to calibrate the spherical heat map automatically. The broadcaster does not adjust the direction of

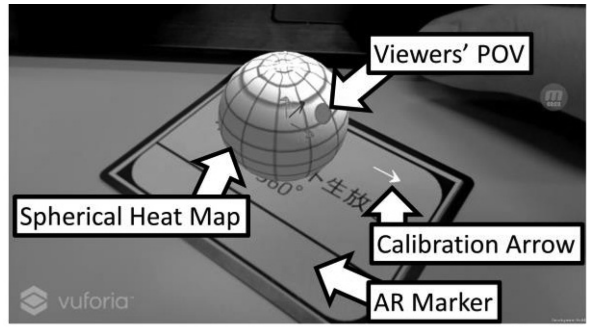

Figure 1: The spherical heat map for the viewers' POV

the omnidirectional camera unlike the conventional web camera because of its characteristic which can shoot a image of 360-degree field of view. For this reason, it is extremely difficult to specify the position of the broadcaster in the 360-degree video. Moreover, it is also necessary to check the position of a display device for the spherical heat map. To solve this issues, we propose a method using an AR marker for the calibration. The AR is a technique of superimposing various virtual objects on the real space and it has a high affinity with the spherical heat map. The AR marker is used to calculate the display position and direction of the AR object. The broadcaster can perform horizontal calibration by rotating the AR marker according to the direction of the omnidirectional camera in the preparation of the broadcasting. After the calibration, it is possible to confirm the viewers' POV without losing the synchronous state by changing the direction of the AR marker following the direction of the omnidirectional camera. To display the spherical heat map, it is preferable to use a portable terminal which can use the AR function and the camera function to display the spherical heat map. We use smartphones and tablets for implementation.

IV. Implementation of Spherical Heat Map

We implemented a prototype system of the spherical heat map for the viewers' POV. Figure 1 shows an example of the spherical heat map in the prototype system and the sphere on the AR marker shows the spherical heat map. The prototype system consists of three components which are a 360-degree Internet live broadcast system, a POV server which collects the viewers' POV information and the spherical heat map. Figure 2 shows the architecture of the prototype system. In this implementation, the spherical heat map was developed as an Android application using Vuforia[6] for an AR library. We use the Ricoh Theta S as an omnidirectional camera for the prototype system.

A broadcaster needs to prepare a PC with an omnidirectional camera for live broadcasting, and a smartphone for displaying a spherical heat map. The broadcaster access a broadcasting system and start 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. The broadcaster will be able to grasp viewers' POV by the spherical heat map launching the smartphone application. The broadcaster can enjoy communication with viewers while grasping the viewers' POV.

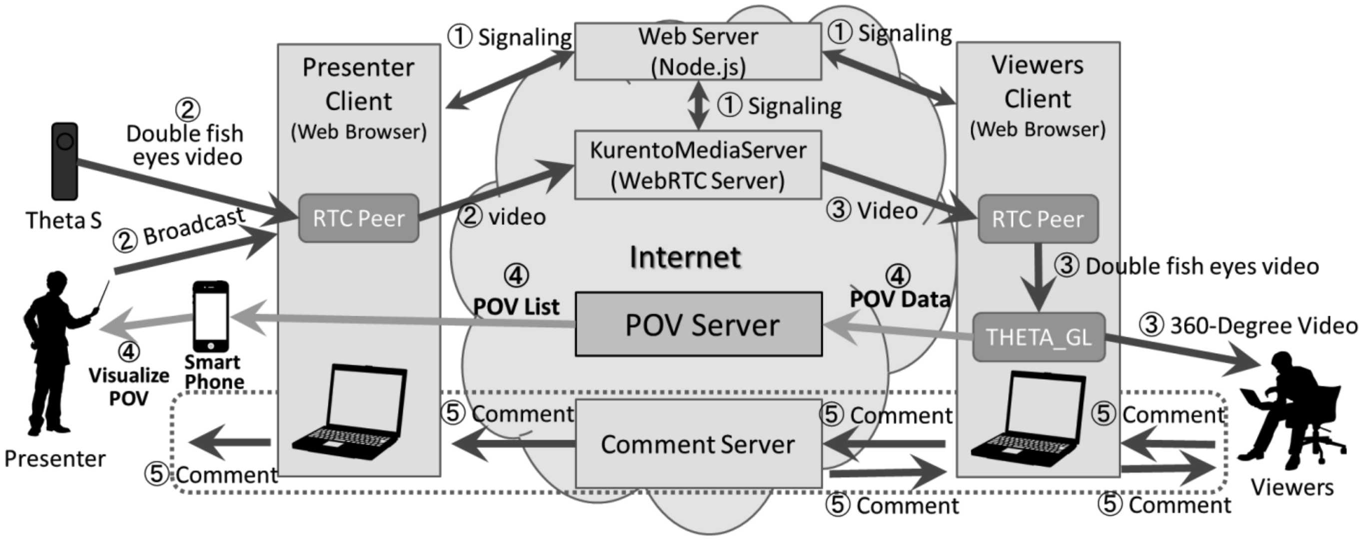

Figure 2: The architecture of the prototype system

Various well-known online video distribution services such as YouTube shift from Adobe Flash to HTML5[7] 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[8] 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[9] 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.

The POV server collects the POV information received from the viewers and converts it into information for creating a heat map. To create the spherical heat map in 3D space, the 3D coordinates, the effect range (radius from the point) and the amount of heat are required. In this implementation, the amount of heat and the effect range are fixed for all POV data. The received POV data consists of a time stamp and coordinate information on the WebGL space. The POV server converts the received data to spherical coordinates of the spherical heat map and keeps the list of the converted data. In order to ensure real-time nature, the coordinate information which has a condition of certain time having passed is deleted from the list. When the POV server receives a request from the spherical heat map, it returns the list of the spherical coordinates.

The spherical heat map creates a heat map based on the received list and visualizes the viewers' POV to the broadcaster. We adopted the shader method[10] for creating the heat map. In this method, transparent small spheres are located in 3D space based on the received list of the spherical coordinates. When an object with shader overlaps one of the located transparent spheres, the heat map is displayed on the surface of the object using a color scale. In order to make the positional relationship of the sphere easier to understand, an image with multiple black lines drawn on a white background were mapped to the sphere. The spherical heat map sends a request to the POV server every second to update the list of the viewers' POV. As the list is updated, it is reflected on the heat map of the spherical surface and the viewers' POV can be visualized in real time.

When implementing the heat map with AR, we need to be careful about how to display the bottom and back of the AR object. The AR system has a characteristic which needs AR markers to be included in the shooting range of the camera in order to display AR objects. Therefore, The AR system is difficult to shoot the bottom of the AR object because the AR marker would be out of the shooting range. The AR marker are used for calibration and cannot be moved. Because of this limitation, the broadcaster needs to move around the AR marker to check the back of the AR Object. The issue is the difficulty to check the whole of the AR object such as spherical heat maps. As a way to solve this problem, we decided to implement a temporary rotation function. This function allows the broadcaster to rotate the AR object in any direction only while performing swipe operation. By allowing the AR object to rotate, the broadcaster can check the viewers' POV even without moving. We also set up three arrows at the center of the spherical heat map to convey the degree of rotation to the broadcaster during the rotation operation. The three arrows indicate the X axis, the Y axis, and the Z axis, respectively. This method is used in 3D modeling software. Broadcasters can check how much the spherical heat map rotates by checking the direction of the three arrows. When the broadcaster finished the swipe operation after checking the viewers' POV, it returns to the original synchronized state by initializing the rotation of the spherical heat map. The broadcaster can confirm the anyplace POV without losing the synchronization state by this function.

We also examined mapping images taken by omnidirectional cameras to the spherical surface of a spherical heat map. However, the mapped video image is reverses horizontally like a mirror when viewed from the broadcaster. The broadcaster is further confused when checking the spherical heat map by using temporary rotation function. Therefore, we decided not to map the image to the spherical heat map.

V. Initial Evaluation

We conducted an initial evaluation of the spherical heat map for the viewers' POV. The participants as the role of the broadcaster actually performed 360-degree Internet live broadcasting using the prototype system and the other participants as the role of the viewers performed some tasks watching the broadcasting. Figure 4 shows the flow of the evaluation experiment. We evaluated the effectiveness of the spherical heat map by comparing the difference between the presence or absence of the spherical heat map. We performed questionnaire and interview surveys after the broadcasting. The experimental results are expected to shorten the time required until the broadcaster responds to comments from the viewers and to improve the accuracy of the response by using the spherical heat map. The experiment was conducted 4 times. 6 different participants attended in each experiment and they were divided into a broadcaster and 5 viewers. The total was 24 participants. The broadcast content is a chat communication between the broadcaster and the viewers about some objects prepared in the room for the broadcaster in advance. The broadcaster and the viewers were allocated in separate rooms. Figure 3 shows the positional relationship between them.

In the broadcasting, the viewer posted some specified comments at some specified timings. The content of the comments are questions about the located objects in the room for the broadcaster by using demonstrative pronouns. We measured the time spent before the broadcaster responded to the comments correctly. In the questionnaire, we asked the broadcaster on a 5-point scale how easy it was to identify objects which was mentioned in the comments. We also asked the viewers how much they felt their comments were understood by the broadcaster on a 5-point scale. Table 1 shows the evaluation results about the broadcaster and Table 2 shows the evaluation results about the viewers. Table 3 shows the results of the response time measurement. The broadcast A in the tables did not use the spherical heat map and the broadcast B in the tables used the spherical heat map.

If the response time to the comment was long, the viewer felt that the comment had been ignored and the satisfaction level of the broadcasting would be worsened. If the response time was short, the viewers could enjoy smooth communication with the broadcaster and the satisfaction level of the broadcasting would be improved. In this experiment, however, the measurement time in the broadcasting with the spherical heat map was increased by approximately 10 seconds comparing with the broadcasting without the spherical heat map. On the other hand, the result of the questionnaire to the broadcaster showed improvement of the communication accuracy. From these results, we consider that the spherical heat map is effective to correctly understand meaning of the comments. In the questionnaire to the viewers, the spherical heat map had also higher evaluation points but significant differences by the t-test could not be confirmed.

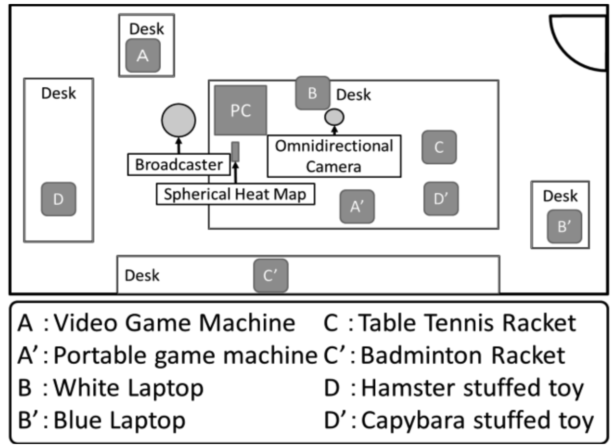

Figure 3: Layout of installed object

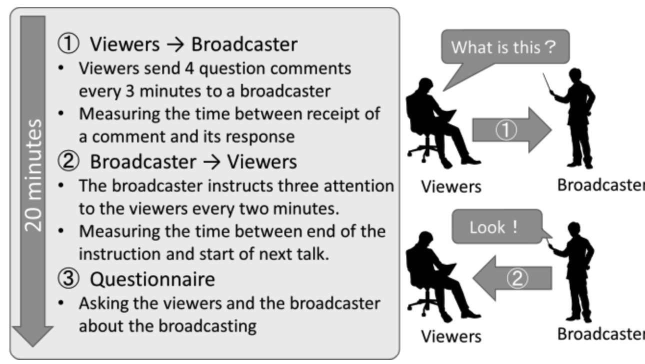

Figure 4: The flow of the evaluation experiment

Table 1: Result of the accuracy of communication on installed objects for broadcasters

Broadcast Type	Object	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average of Object	Average of Broadcast
Broadcast A (absence of system)	A	1	3	1	4	2.25	2.8125
	B	2	3	2	5	3	
	C	1	3	3	5	3	
	D	3	2	2	5	3	
Broadcast B (presence of system)	A'	1	5	4	5	3.75	3.5625
	B'	4	3	4	4	3.75	
	C'	4	2	5	4	3.75	
	D'	3	3	3	3	3	

Table 2: Result of the accuracy of communication on installed objects for viewers

Broadcast Type	Object	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average of Object	Average of Broadcast
Broadcast A (absence of system)	A	3.8	3.2	3.2	3.0	3.3	3.6625
	B	3.4	4.2	4.0	3.6	3.8	
	C	4.0	3.8	3.6	3.8	3.8	
	D	3	4.2	4.0	3.8	3.75	
Broadcast B (presence of system)	A'	3.8	4.2	3.4	2.4	3.45	3.5625
	B'	4.4	3.8	3.8	5.0	4.25	
	C'	4.0	4.4	4.8	3.4	4.15	
	D'	3.6	4.6	3.2	3.4	3.7	

Table 3： Results of response time

Broadcast Type	Target	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Average of Experiment	Average of Broadcast
Broadcast A (absence of system)	Object A	21 Sec	22 Sec	6 Sec	24 Sec	18.25 Sec	20.875 Sec
	Object B	9 Sec	39 Sec	13 Sec	26 Sec	21.75 Sec	
	Object C	43 Sec	33 Sec	12 Sec	19 Sec	26.75 Sec	
	Object D	4 Sec	24 Sec	19 Sec	20 Sec	16.75 Sec	
Broadcast B (presence of system)	Object A′	45 Sec	39 Sec	46 Sec	38 Sec	42 Sec	30.875 Sec
	Object B′	29 Sec	26 Sec	29 Sec	16 Sec	25 Sec	
	Object C′	35 Sec	32 Sec	24 Sec	32 Sec	30.75 Sec	
	Object D′	29 Sec	33 Sec	13 Sec	28 Sec	25.75 Sec	

From the experiment, we confirmed that the spherical heat map for the viewers’ POV was effective to improve communication accuracy. The reason why the significant difference could not be confirmed was that the operation to look the spherical heat map might be burdensome for the broadcaster. We believe that this point can be solved if we can shorten the time required for looking the spherical heat map by improving the user interface. By shortening the time, effectiveness of the spherical heat map could be confirmed clearly in both the communication accuracy and the response time.

VI. DISCUSSION

As the result of the evaluation experiment by using the spherical heat map, we found that the broadcaster would be able to understand the viewers’ comments more correctly. Although the broadcasting without the spherical heat map had a short response time, there was a case where the broadcaster misunderstood the viewer 's question and talked about the different object. On the other hand, in the broadcasting with the spherical heat map, there was no case where the broadcaster misunderstood the viewer 's question a although the response time increased. From this point of view, we believe that the spherical heat map is one of the effective approaches to support communication in the 360-degree Internet live broadcasting using an omnidirectional camera.

In the experiment, however, the omnidirectional camera was fixed in the room. It was not changed the position and the direction. For the Internet live broadcasting in the mobile environment, further functional improvements will be required. For example, it is difficult to carry the AR marker. Since the direction of the omnidirectional camera will be frequently changed in the mobile environment, it is difficult to synchronize the heat map with the direction of the 360-degree image. In that case, the broadcaster needs to check the heat map while considering the difference.

These problems have in common that it is caused by the overhead when the broadcaster check the spherical heat map. Therefore, it is necessary to develop an assistant function to analyze and summarize the viewers’ POV to give advices to the broadcaster. The assistant function analyzes the viewers’ POV by the cluster analysis and tells the broadcaster where is the center of interests. Moreover, it will be possible to emphasize important information by estimating the degree of interest for the broadcast based on the movement of viewers' POV. As a result, the broadcaster will be able to grasp the state of the viewers without checking of the heat map by themselves.

VII. CONCLUSION

We pointed out a problem of communication in the 360-degree Internet live broadcasting under the importance of gaze information in communication. As a solution to the problem, we proposed a viewers’ POV visualization method using a spherical heat map in the 360-degree Internet live broadcasting. The spherical heat map visualizes density of the viewers’ POV on the sphere handling the POV data as vector data from the center point of the sphere. We implemented the prototype of the proposed model and also made an initial evaluation. As the result, we found a trend that the spherical heat map improved the accuracy of the communication between the broadcaster and the viewers.

In the future work, it is necessary to improve the user interface of the spherical heat map and the method to notify the broadcaster of the viewers’ POV. In addition, new products of omnidirectional cameras are announced one after another and we need to conduct a further study about their usage in various locations. In order to reduce the overhead the broadcaster, we will study some assistant functions which analyze viewers' POV instead of the broadcaster.

REFERENCES

[1] Roel 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] David M. Grayson, Andrew 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] Naoki Mukawa, Tsugumi Oka, Kumiko Arai, Masahide Yuasa, “What is connected by mutual gaze?: user's behavior in video-mediated communication”, CHI EA '05 CHI '05 Extended Abstracts on Human Factors in Computing Systems, pp. 1677-1680, 2005

[4] Keisuke Shiro, Atsushi Okada, Takashi Miyaki, Jun 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

[5] S. Rothe, et.al, ``CVR-Analyzer: A Tool for Analyzing Cinematic Virtual Reality Viewing Patterns,’’ in Proc. of ACM MUM 2018, pp. 127-137, Nov. 2018.

[6] Vuforia Developer Portal, https://developer.vuforia.com/, November 7, 2018.

[7] YouTube Creator Blog: Hot and Cold: Heatmaps in VR, https://youtube-creators.googleblog.com/2017/06/hot-and-cold-heatmaps-in-vr.html, June 16, 2017

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

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

[10] Arrays & shaders: heatmaps in Unity - Alan Zucconi, https://www.alanzucconi.com/2016/01/27/arrays-shaders-heatmaps-in-unity3d/, January 27, 2016