What are we looking at? An interactive 360° Video Viewing Experience
By Yi Xie

Introduction

While Virtual Reality and 360-degree video are getting more attention from individual users and entertainment industry, how these immersive viewing environments can enhance the narration for the content creators has not been fully explored yet. The creators are given the audience more freedom in this omnidirectional world since there is not a particular perspective. However, understanding the audience's viewing experience and behavior in this omnidirectional world is crucial for making this visual experience more effective and unique. My research aims to develop a 360° video streaming environment that tracks and analyzes audience focus and navigation habit and also gives the future viewers access to these data as a viewing guide.

In recent years, video service providers such as YouTube and Facebook have invested significant amount of resources to develop a dynamic streaming environment for virtual reality. According to Facebook newsroom, more than 20,000 360º videos have been uploaded to Facebook and hundreds more of them are added every single day[1]. Recently, Facebook added two new features to enhance the 360 viewing experience[2]. Guide lets 360º video makers to set a default path by selecting points in the video that they want people to see in order to create a lean-back watching experience. The viewing window will automatically move around the 360º video sphere and focus on the selected spots so that the audience doesn't miss important things while watching the video. Another feature called heatmap allows the content creators to know where people are looking inside the 360º video by visualizing the most viewed angles whenever viewers move more than 30 degrees. This feature helps the creators to better understand which part of the video is more eye catching.

However, the data they collect is only available to the creators and only tells them the most viewed angles in their video. Many previous researches on 360º audience behavior were also limited as their testing videos were relatively short in order to avoid any discomfort from using headsets[3]. These short videos usually have a loose or non narrative structure. 360º documentaries or fiction films with a clear narrative haven't been explored in great details. Many questions regarding navigation patterns, user's' interaction, and content of the video remain open. For instance: how often do people change their viewpoints while watching 360º video? Does viewer's behavior correlate to the content or the duration of the video? Is it helpful for the first time viewers to know what previous audience were focusing on? How

[1] New Steps Toward the Future of Virtual Reality. Feb. 21st, 2016. https://newsroom.fb.com/news/2016/02/new-steps-toward-the-future-of-virtual-reality/
[2] Evgeny Kuzyakov, Shannon Chen, Renbin Peng. Enhancing High-Resolution 360 Streaming with View Prediction. Apr. 19th, 2016. https://code.facebook.com/posts/118926451990297/enhancing-high-resolution-360-streaming-with-view-prediction/
[3] Xavier Corbillon, Francesca De Simone, and Gwendal Simon. 360-Degree Video Head Movement Dataset. In Multimedia System Conference (MMSys), 2017.

would these audience behavior data benefit the content creators? Can they provide extra ancillary information based on the audience's feedback?

This research will run 360º viewing experiments to collect audience behavior dataset and visually display these information through the streaming environment, making the data accessible to the audience and fostering a more engaging social watching experience. We also expect to analyze the relation between the content and the audience reaction and hopefully conclude with some pattern to help the content creators to improve or expand on their visual project.

Related Work

In the past, a lot of audience behavior analysis have been done on 360º images. De Abreu et al. perform a study of navigation patterns for 360º image viewing sessions on a Head-Mounted Display (HMD) and make the dataset available to the general public.

Corbillon et al make a step further and introduce an open-source software that allows anyone to produce dataset of 360º video viewing session. They also include user study with 59 users watching five 70 second long 360º video on a HMD and perform some brief content-driven analysis.

However, most researches that have been done so far uses HMD, which is not the most common way for user to consume 360º videos, especially for the longer ones with a form of narration or plot. The data analysis is also heavily content-driven because of the short length of the video samples. This paper is going to focus on longer 360º video contents displayed on a flat screen and analyze audience behavior in correlation to content of the videos.

Method

In this section, we describe the software implemented to capture user's movements data during 360º video with a computer screen, as well as the test material and test subjects during the viewing session.

1. Software

We developed a 360-degree video streaming environment to capture and save a log file to record each user's viewing position at each frame or whenever a mouse movement occurs during the visualization of the 360-degree video on the computer screen. We decide to present the video content on a computer screen instead of other VR devices because it is the most common and accessible way of viewing 360 content nowadays. So the user will hopefully be more comfortable using the technology and react naturally closer to their real life viewing experience. The main purpose of the streaming environment is to accurately record any view changes to a timestamp corresponding to a video timestamp at frame level. Thus, the user behavior will be able to analyzed in relation to the visual changes happen in the visual content.

Each log file of the user's behavior data has the same structure and saved as a csv file. There is one sample per line. Values are separated by columns, according to the following format. Table 1 is an example of part of a user log file. Column A records each frame number where there is a mouse movement. Column B represents the updated x-coordinates of camera angle at a specific frame. Column C represents the updated y-coordinates of camera angle at a specific frame.

	A	B	C
1	402	-0. 0842478	0
2	404	-0. 1684956	0
3	405	-0. 9266761	0. 1684845
4	406	-2. 27473	0. 7582582
5	407	-4. 717716	2. 190353
6	408	-7. 834783	3. 706764
7	409	-11. 2047	5. 307473
8	410	-15. 16421	7. 076619
9	411	-19. 46022	8. 508619
10	412	-23. 67289	9. 688167
11	413	-27. 37991	10. 69917
12	414	-32. 01339	11. 8786
13	415	-34. 20369	12. 13133
14	416	-35. 88845	12. 55252
15	417	-37. 23648	12. 63677
16	418	-38. 41595	12. 63677
17	419	-39. 51117	12. 63677
18	420	-40. 69056	12. 55253
19	421	-41. 95426	12. 38404
20	422	-44. 14445	11. 6259

Table 1: Example of part of a user log file.

The software has been developed in C# in cross-platform game engine Unity. The streaming environment is based on the UMP, "a universal media framework plugin for Unity that based on Video LAN Codec (libVLC) native libraries". We performed all user studies and the data collection on Windows 10 Operating System, but the software should be compatible with any Windows, Mac, and Linux OS. The users use mouse to navigate through the 360-degree environment and the audio is played through an output headphones (Audio-Technica).

2. Test Material

The navigation patterns in the dataset have been collected on two 360-degree YouTube videos, described in Table 2 below. We limited the number of test materials to be able to do each user study session in a reasonable duration so that the user can perform more similarly to their normal watching experience and also to encourage more participants. We chose two relatively long 360-degree videos both with a narrative purpose in order to receive any meaningful result associates to their narration. The two videos belong to two completely different genres, one computer-generated 360-degree animation, and the other realistic documentary shot with 360-degree camera. The reason why we chose these two genres is because 360º technology are most widely used to produce animation or documentary. We assume the viewer will respond different between different genres of 360-degree content, so the contrasting video content has been chosen to study any potential navigation patterns in relation to the genre of the video.

Prior to the official viewing experiment, the participant will first watch a short 360-degree training video from YouTube in order to get them comfortable using the mouse and familiarise navigating in the streaming environment. The video content includes directed arrows to direct them to find certain important elements in the video and to encourage the viewer explore the entire 360-degree environment. The rotation speed of the mouse may be altered upon user's request after finishing the training video.

Each video file has been downloaded in equirectangular format, at the maximum resolution and bit-rate available on YouTube.

3. Viewing Session

All participants to our user study performed one viewing session of a total duration of 14 minutes. Before the beginning of the session, a consent form explains the main steps of the viewing session and any potential viewing discomfort is presented to the participant. Each user is also informed that their navigation pattern will be recorded and need to voluntarily sign the consent form to continue the session.

After signing the consent form, the participant is asked to fill out a brief survey on a piece of paper, concerning the user's gender, age, and level of familiarity with computer, video games and 360-degree content.

Then the user will be asked to put on the headphones. The training session starts with a one-minute-long documentary video about pandas. The viewers were orally instructed to move their mouse in order to familiarise with the 360-degree viewing environment. The rotation speed of the mouse may be altered upon user's request after finishing the training video.

The two testing videos will then played consecutively in a randomized order.

At the end of the viewing session, the user is asked to remove their headphones. The user is then given access to a variety of refreshments. A questionnaire regarding their viewing experience to each specific 360-degree videos needs to be filled out after the viewing session.

4. User Sample

36 users took part to our user study in November, 2017. Most people in the sample are students or staffs member from Connecticut College in New London, CT. For each user, the data is collected in a dedicated csv file named as "Date_Number of test within that day" (i.e 20171110_01.csv). Two viewers stopped during the viewing session due to physical discomfort caused by the motion sickness and their files are deleted and not considered for the analysis.

More detailed statistics about this sample group is listed below:

- Number: 36 total participants (2 are not usable)
- Gender: 11 males (32.4%) and 23 females (67.6%)
- Age range: 18 - 64 year old
- Average time on computer: 5.7 hrs per day
- Past Experience with 360º Video:

- No Experience: 16 (47.1%)
- Watched 360º Videos: 17 (50%)
- Created 360º Content: 1 (2.9%)

<u>Audience Viewing Pattern Analysis</u>

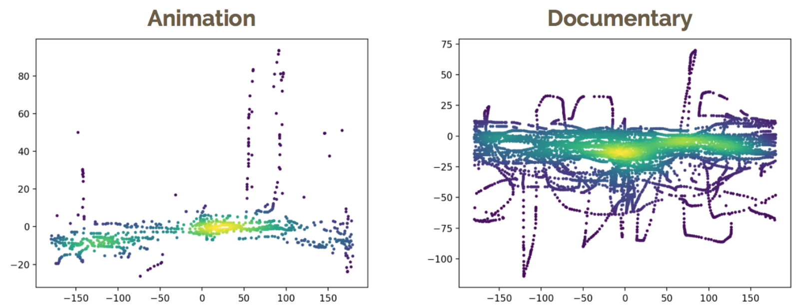

Figure 1: A scatter plot that visualize the user's movement in the viewing session.

To better understand viewer's viewing pattern, we developed a data visualization tool using Javascript C3 Library. The navigation pattern is drawn on a scatter plot and the color intensity represents the point of view distribution in the 360º world. An example is shown above in Figure 1. The visualization tool can also concatenate multiple users' data and draw them on one single scatter plot. It is also capable of display segments of data by selecting the range of frame counts in the video.

The first step of the data analysis is to understand how active the user is. The active rate of each user in each video is calculated using this formula: active rate = number of frames that the user moves the mouse/ total frame count in the video. From the figure 2, it is clear that users are more active in documentary (30.2%) than in animation (20.6%). In addition, we study the active rate in relation to user's familiarity with the 360º technology. The result shows more experienced viewer is less active in both films and viewer with no previous experience are the most active group.

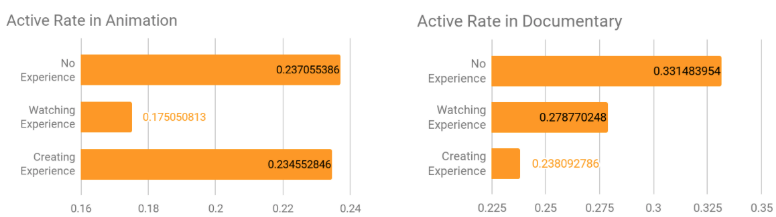

Figure 2: User active rate in animation and documentary

Another common viewing behavior that we discover in both animation and documentary is that there is a decrease in interest when the same object/character appears multiple times in both genres. The more times the object/character is mentioned, the viewer is less likely to follow the movement with it.

In addition, there are several other interesting findings in the documentary and animation. User tends to follow and focus on character's movement more often in animation than documentary. This result can be a signifier of different level of identification. In animation, the little girl is the only character of the movie and the audience is forced to focus and identify with this protagonist. On the contrary, multiple characters are introduced throughout the film and the narrative structure is pretty loose compare to the animation. We believe that the narrative structure and character setting directly influence the level of identification within the film. Therefore, users are more likely to follow in the animation than the documentary.

The last discovery is the relation between the shot size of each scene and the active rate in documentary. After breaking the documentary down to each scene and calculate the active rate in each scene, the result suggests that users are 5% more active in wider long outdoor scenes than tighter medium/close-up shots indoor. This behavior can also be explained using film theory. The use of smaller shot size forces the audience to focus and identify with one specific object/character on screen and sometimes will cause physical response and feeling such as claustrophobia and fear. Wide shots are more commonly used to establish entire environment/atmosphere. Therefore, based on the established camera language in traditional cinema, when a wide shot is shown in 360º degrees, viewers are more willing and comfortable to navigate and look around.

Audience Viewing Data Visualization

Apart from the analysis of user viewing behavior, the 360º streaming platform also aims to use effective visual representations to display past viewing data to future audience. The past viewer's viewing positions are visualized on an x-axis bar at the bottom. While the video is playing, the user now can see the most popular viewing spots in the video at the given moment (Figure 3 below). The entire 360º degree in horizontal direction is represented in the 36 blocks at the bottom. Each block equals to 10 degrees. Audience's current position is updated in real-time using the yellow block. When there are viewer within the 10-degree block, a transparency value of the green color is assigned. The most populated angles will have lower transparency, and therefore, the green will be more obvious. By making the data accessible to future audience, this 360º streaming environment aims to provide a visual guidance for future navigation and to foster a more engaging social watching experience.

Figure 3: Screenshot of 360º Streaming Platform with visual guidance in Unity

Conclusion

In this paper we presented a dataset including viewing positions of 34 participants recorded while they were watching two 360º films of different genres, animation and documentary. The user data is visualized and analyzed focusing on how the film genre and content affect the viewing behavior. Some meaningful results and conclusions were drawn referencing some film theory. We expect this analysis will help filmmakers to understand 360º video consumption. Based on the data we received, we add another feature in the 360º streaming platform to visualize past viewer's data to assist future consumption. This updated platform will hopefully maximize communicativeness and accessibility of 360º content to the viewer and foster a more engaging social watching experience to the audience.

References

Ana Kuzmanic Skelin, Tamara Grujic Supuk, and Mirjana Bonkovic. Visual Peoplemeter: A Vision-based Television Audience Measurement System. Advances in Electrical and Computer Engineering: Vol. 14, No.4, 2014.

Evgeny Kuzyakov, Shannon Chen, Renbin Peng. Enhancing High-Resolution 360 Streaming with View Prediction. Apr. 19th, 2016. https://code.facebook.com/posts/118926451990297/enhancing-high-resolution-360-streaming-with-view-prediction/

John J. Hussar. 360 Degree Spherical Video: The Complete Guide to 360-Degree Video. New York: Grey Goose Graphics, LLC, 2016.

New Steps Toward the Future of Virtual Reality. Feb. 21st, 2016. https://newsroom.fb.com/news/2016/02/new-steps-toward-the-future-of-virtual-reality/

Xavier Corbillon, A Devlic, G. Simon, and J. Chakareski. Viewport-adaptive Navigable 360-degree Video Delivery. In IEEE ICC, 2017.

Xavier Corbillon, Francesca De Simone, and Gwendal Simon. 360-Degree Video Head Movement Dataset. In Multimedia System Conference (MMSys), 2017.