Guiding the Viewer in Cinematic Virtual Reality by Diegetic Cues

Sylvia Rothe[1] and Heinrich Hußmann[1]

[1] LMU Munich, Germany
{sylvia.rothe, hussmann}@ifi.lmu.de

Abstract. Cinematic Virtual Reality has been increasing in popularity the last years. Watching 360° movies with a Virtual Reality device, viewers can freely choose the viewing direction, and thus the visible section of the movie. In order to ensure that the viewer observes all important details, we investigated three methods of implicitly guiding the attention of the viewer: lights, movements, and sounds. We developed a measurement technique to obtain heat maps of viewing direction and applied statistical analysis methods for spatial data. The results of our work show that the attention of the viewer can be directed by sound and movements. New sound induces the viewer to search for the source of the sound, even if not all participants paid attention to the direction of the sound. In our experiments, lights without movements did not draw more attention than other objects. However, a moving light cone changed the viewing direction considerably.

Keywords: Cinematic Virtual Reality; 360° movie; guiding attention; spatial sound; directing gaze.

1 INTRODUCTION

360° movies are attracting widespread interest and have many possible applications, e.g. telling stories about exciting locations in the world or ancient places of interest in history. Especially, museums and other educational institutions can take advantage of this.

In Cinematic Virtual Reality (**CVR**) the viewer watches 360° movie using a Head Mounted Display (**HMD**) or other VR devices. Thus, the viewer is inside the scene, and can freely choose the direction of view. Accordingly, the viewer determines the visible section of the movie – the field of view (**FoV**). Therefore, it is not always possible to show the viewer what is important for the story. Several conventional filmmaking methods for guiding the viewer's line - such as close ups or zooms - are not practicable in CVR. For other methods, it needs a closer analysis of whether they are suitable to direct the attention of the viewer to important details in a CVR environment. In this paper, we focus on some traditional alertness methods of filmmaking which we think are transferable to CVR: sound, lighting, and movements.

From film theory, we absorb the terms **diegetic** and **non-diegetic**. Diegetic cues are part of the scene – for example, a musician playing music. Non-diegetic cues come from outside – for example, film music or a voice over. The cues considered in this paper are diegetic cues.

We want to investigate if diegetic cues are suitable for drawing the attention of the viewer. For this the following questions should be studied:

- Which **type** of cues are more effective for guiding the attention: sound, light or movements?
- Which **combination** of cues sound/light, light/movement, movement/sound are efficacious for guiding the attention?
- Can a diegetic **audio** cue draw the attention of the viewer to it, even if it not in the field of view in the moment the cue appears? How important is the direction of the sound for guiding?
- Can a diegetic **moving** cue draw the attention of the viewer to it, even if it not in the field of view all the time?

For answering these questions, a short movie was produced, which was shown to 27 participants. The head tracking data were recorded for generating heatmaps and analysing the data. For finding significant hotspots spatial statistics methods were used.

2 Related Work

360° movies are not new and not only produced for HMD. Investigations for 360° videos on a desktop [1] and in full domes [2,3] provide general background information.

Much research in recent years has focused on **presence** in VR environments [4,5,6,7] – these results can be adapted to CVR. Poeschl et al. [8] and Serafin [9] have shown that spatial sound is important for a high level of presence in virtual environments. In our study, we concentrated on guiding the attention of the viewer. However, we considered the results of these investigations and endeavoured to ensure that our methods do not interfere with the presence.

Syrett et al. [10] have discovered that some viewers feel distracted by the freedom to choose the viewing direction. In their experiments, it happened that important parts of the storyline were missed. In the literature [11,12,13] several methods for guiding the viewer are explored for non-VR environments, such as salient objects, sounds, lights, or moving cues. Our work examines how this can be adapted to CVR, even if an object is not in the FoV of the viewer.

Van der Burg et al. [14] showed that audio cues (pop) synchronized to a salient visual cue (pip) reduces the search time, even if the audio cue does not have any location information. Emil R. Høeg et al. [15] enhanced this experiment to Virtual Reality with sound cues from the same direction as the visual cue. They demonstrated that binaural cues lead to shorter search times, even though the visual cue was not always visible at the moment the audio cue was presented. In the experiments, the participants were given a search task in an abstract VR environment. In our study in comparison to [14,15], we move closer to a real cinematographic setting by using a realistic scene instead of abstract symbols and by not giving a concrete task to the participants but letting them choose freely what to do next.

Several investigations are focused on **non-diegetic** methods. Lin, et al. [16] compared two focus-assistance tools. The first one, an autopilot, adjusts the field of view automatically, so the viewer can see the target. The second tool uses visual signs that show the direction which the viewer should follow. Both tools assist the viewer to find an intended target. Questionnaires were used for investigating ease of focusing, presence, and discomfort. The results illustrate that the preferred method depends on the viewer's preferences and the content of the movie.

Brown, Cullen, et al. [17] describe several gaze attraction techniques for VR which are also useable for CVR: two guided camera techniques and two voluntary distractor methods. The first technique is a scene transition by fading out the previous scene and fading in the subsequent scene with the important section in the FoV. The second method is a forced camera rotation, similar to the autopilot of Lin, et al. [16]. The authors expect disruption and disorientation of the viewer through this method. Additionally, two distractor techniques are described. First, a firefly drifts in the FoV until the viewer changes the viewing direction to the target following the firefly. Secondly, spatial sound is used for guiding the attention. The investigations are not finished yet and the results not published.

Few researchers have addressed how to guide the viewer's attention in CVR by **diegetic** cues. In 1996 Pausch et al. [18] examined how the attention of the viewer can be drawn to a desired spot. For that, they used the characters as diegetic cues. Similarly, Sheikh et al. [19] connected several diegetic cues (motion, gestural, and audio cues) to the main character of a scene. In the experiment, the cues with an audio component were more helpful than just visual cues, even if the sound was not fully spatialized. However, the main character of a movie attracts more attention, in general, even without any special cues. In our approach, we investigate if it is possible to guide the view with cues connected to neutral objects that the viewer has not seen before.

Nielsen et al. [20] compared a diegetic cue (firefly) with a non-diegetic cue (forced rotation) and no guidance. Using a questionnaire, they figured out that the diegetic cue (firefly) was more helpful than the non-diegetic cue (forced rotation). Furthermore, the results demonstrate that the non-diegetic cue may decrease the presence.

For determining which cues attract the attention of the viewer and can change the viewing direction, we decided not to use questionnaires. Instead, the head direction was recorded and evaluated to obtain more precise results. For this, we developed a tool that generates heatmaps for every timecode in the movie.

3 Methods

Material

To investigate whether diegetic cues can guide the attention of the viewer, we produced a movie which contains various diegetic cues

- **s**ound from a certain direction (**s**)
- **l**ighted objects (**l**)
- **m**ovements of stationary objects – for example swinging (**m**)
- **lo**comotive objects - movements with change of position (**lo**)

In the scenes which are relevant for the test, the sound is spatial and connected to visual objects. In the scenes which are used for randomizing the viewing direction, the sound is not spatial.

The field of view of every viewer was recorded, so it was possible to evaluate if the view was influenced by the cues. For every timecode of the movie a heatmap was generated.

Participants

27 Participants (10 women and 17 men) watched the movie with a head-mounted display (Samsung Galaxy S6 with google cardboard) and headphones (Ultrasone HFI 780). The voluntary participants were in the age between 16 and 73 (mean=40). One participant was a sound engineer and was removed from the data set. He explored how spatial sound worked in the test environment by unusual head movements. In the concluding interview, we noticed that his attention was not on the movie.

We chose a within-subject test design – all participants watched the same movie. Standing they turned head and body to watch the movie. There was no special task, they should look around and follow the objects they are interested in.

After watching the movie, a short interview was conducted. The participants enumerated the objects which they could remember spontaneously. Afterward, they were asked whether they have seen the other objects or not. Our research did not investigate the influence on cognitive assimilation, however this additional information helped to interpret the measured data.

Movie

The movie consists of 4 sections. The sections are separated by a neutral scene: a forest which looks similar in every direction (Fig.1), the sound is not spatial. This scene makes sure that the viewing direction at the beginning of the next scene is random.

Fig. 1. The forest separates the tests scenes to make the viewing direction for the start of the next scene random.

All the other scenes are of the same place - a mystery kitchen of an old castle (Fig. 2). Objects with different cues appear and disappear from time to time. The sound in these scenes is spatial and connected to the appearing objects. We used Unity for attaching the sounds to the objects in the movie.

Fig. 2. The place of the test scenes

In the first part of the movie (scene 1 and 2) the objects do not change their positions. Objects with movements are swinging or flickering. Other objects are lighted or connected with sound. All the objects are there from the start to the end of the scene. It was investigated if these objects can attract the attention of the viewer.

The aim of the second part (scene 3 and 4) is to evaluate if we can modify the viewing direction of the participant by objects, which are not necessarily in the FoV at the beginning. In scene 3 objects are associated with sounds. In scene 4 we use locomotive objects.

In the **first scene**, the investigated objects are connected to only one cue. There is one moving object (m), one object with spatial sound (s), and one lighted (l) object in the room.

In the **second scene**, every object is provided with two cues:

- a lighted moving object (m,l),
- a moving object with spatial sound (s,m),
- a lighted object with spatial sound (l,s).

Thus, all combinations of the three methods are in the scene. Fig. 7 shows the arrangement of the objects.

In the **third scene,** we explored if it is possible to change the direction of view using objects with sounds. For this, several objects with and without sound appear and disappear in the kitchen. At the beginning, there is a ticking alarm clock followed by a whistling tea kettle and a cuckoo clock at different positions. Subsequently, two phones – only one with sound - appear. After they have disappeared again, a gramophone (with sound) and an old record player (without sound) are visible in the kitchen – succeeded by an old radio. The objects are positioned at several positions (Fig. 3) and not always in the FoV of the participants. All the sounds are spatial.

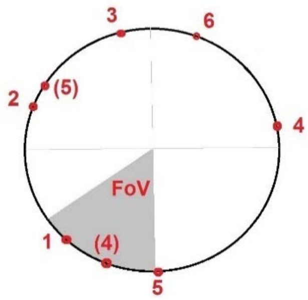

Fig. 3. The Arrangement of the objects in the room, the shaded segment shows the size of FoV (60°), the numbers show the sequence of appearing: 1-alarm clock, 2-kettle, 3- cuckoo clock, 4-phone (4)-silent phone, 5-gramophone (5)-silent record player, 6-radio

The **fourth scene** explores locomotive objects changing their position (lo) with and without additional cues. We tested three different methods for guiding the view to another direction:

- locomotive object without sound – feather (lo)
- locomotive light without sound - light cone (lo, l)
- locomotive object with spatial sound – bee (lo, s)

Statistical evaluation

In a first step, we inspected the heatmaps for relevant timecodes in the movie. We found several hot spots which had to be verified for significance. Therefore, we used spatial statistical methods.

The collected data are point incident data. Point incident data are points connected to an event – in our case the viewer looked to this point. We were interested in significant clusters. To find such clusters, we used the Getis-Ord Gi* statistic [21] – a spatial statistic method for examining spatial data. This statistical method requires values for the investigated points. In order to use this method, the incident data were aggregated and incident counts established. The incident counts – in our case the number of views - are the attribute values which are analyzed by the method.

The Getis-Ord Gi* statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}}$$

where x_j is the attribute value for point j, $w_{i,j}$ is the spatial weight between point i and j, n is equal to the total number of points and:

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2}$$

To apply this statistical method, the GIS software ArcGIS Pro was used. For every point in the incident data set, the p value was calculated and can be displayed by double-clicking on the point. The p value is the probability that the observed pattern was created randomly. A small p value means that the pattern is most likely caused by a cluster. Segments with p values smaller than 0.01, which means 99% confidence, are displayed in red. Fig. 4 shows the color for the different confidence levels. These legend is used for interpreting the results in the next chapter.

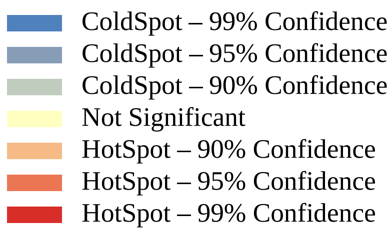

Fig. 4. Legend for the significance test of the hotspots

4 Results

Scene 1

In the first scene three objects with one cue were examined: socket (lighted), ladle (moving/swinging), dripping water (sound). Fig. 5 demonstrates the locations of the objects in this scene.

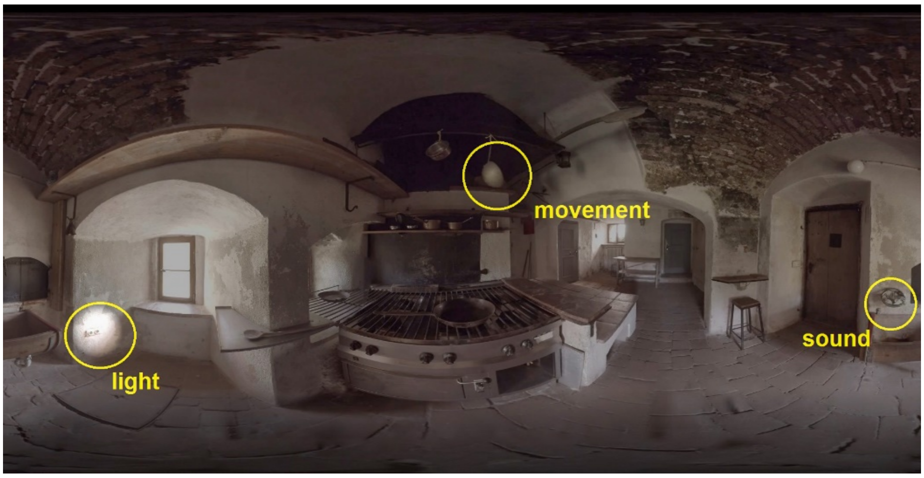

Fig. 5. Objects with one feature: socket - light (l), ladle - movement (m), water – sound (s)

We could not observe any greater attention to one of the cued objects. As can be seen in Fig. 6 the hot spots were distributed over the whole room, except the bottom and the ceiling.

The significance test (Fig. 6, right) showed limited hotspots. The sound object and the moving object each have a significant cluster. At the position of the light (not moving), there was no significant cluster. Instead there were clusters on other objects.

The scene takes 17s (0:20-0:37). In the interview, socket and water were listed by only 3 persons. The ladle was swinging until the end of the movie and enumerated by 16 participants. In a detailed review of the data, we saw that most people inspected the ladle later for a longer time, not in the first scene.

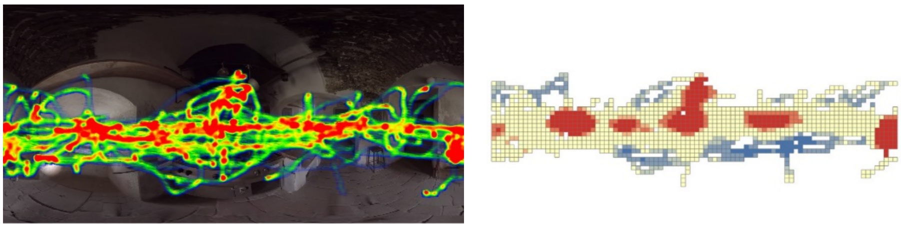

Fig. 6. In scene 1 many little hotspots were generated

Scene 2

The second scene has three cued objects: a flickering lamp (m,l), a pot on the fire (s,m), and a ticking clock (l,s). The arrangement of the objects is displayed in Fig. 7.

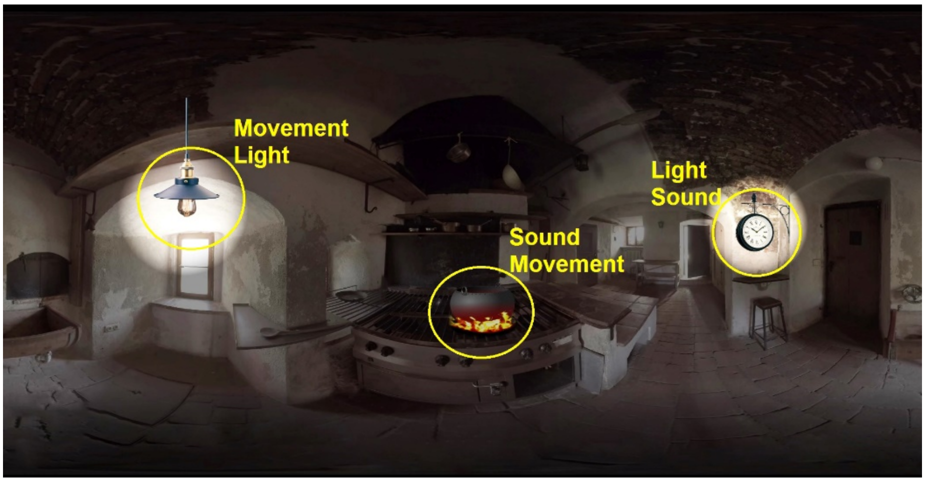

Fig. 7. Objects with two cues: lamp (m,l), pot (s,m), clock (l,s).

As can be seen in Fig. 8, two significant hotspots arose around the objects with sounds (fire, clock). The lighted flickering lamp (without sound) was not recognized by the participants. In addition to the absent sound this could be due to the position of the lamp. It is located higher than the other objects.

The scene takes 18s (0:49-1:07). No participant mentioned the lamp, the fire was listed by 18 people and the clock by 12.

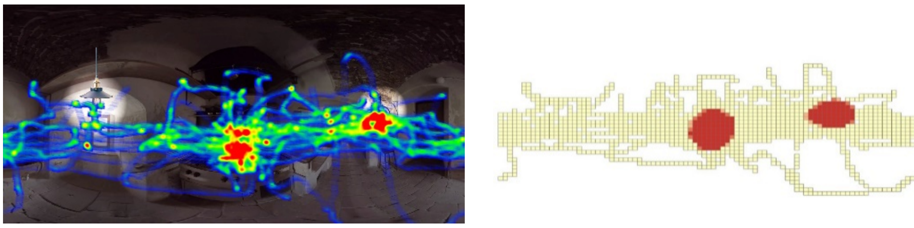

Fig. 8. There are hotspots on the fire and the clock, but no hotspot on the lamp.

Investigating the beginning of the scene (0:50-0:52) we discovered that most participants first looked to the fire.

Scene 3

In the third scene, several objects with spatial sounds appear and disappear. It starts with a ticking alarm clock, 10s later added by a whistling tea kettle and after another 10s by a cuckoo clock, all at different positions (Fig. 9).

Fig. 9. Position of the first three sound objects in Scene 3.

For each of these objects, most participants changed the viewing direction when the objects with sound appeared. As can be seen in Fig. 10 and 11 significant hotspots were generated in all three cases.

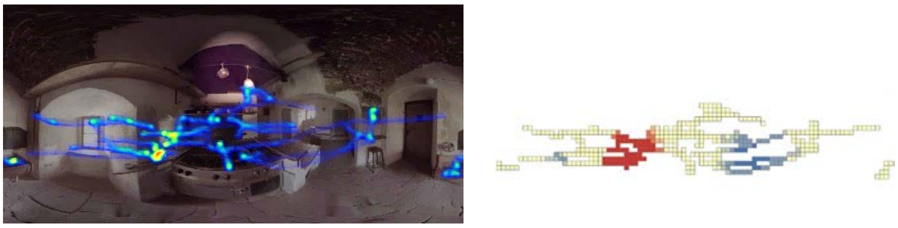

Fig. 10. Heatmap and hotspot for the alarm clock

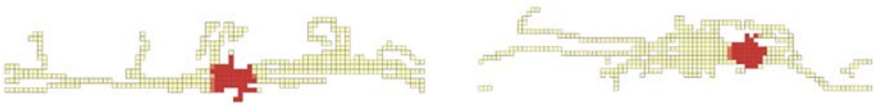

Fig. 11. Left: hotspot for the whistling kettle, right: hotspot for the cuckoo clock

Afterwards, two phones – only one with sound – appear (Fig. 12). Around both objects a hotspot has formed even though only one phone was ringing (Fig.13).

Fig. 12. Two objects: phone with sound (right), phone without sound (left)

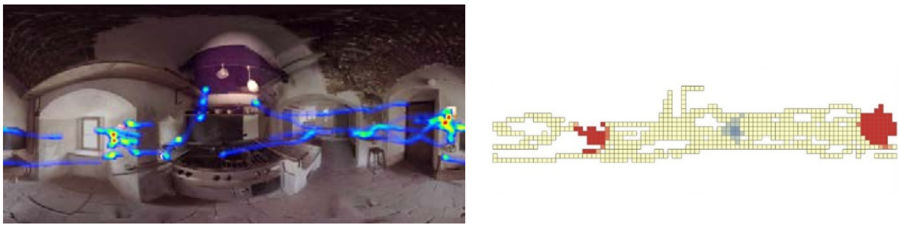

Fig. 13. Heatmap and hotspots for the phones with and without sound

Analysing the data, we found out that the phone with sound has a stronger effect. In the interview, more people listed the ringing phone near the door, (11 participants) than the phone without sound near the window (8 participants). Seven participants listed both phones.

In the next step, this experiment was repeated with other objects: a gramophone (playing an old melody) and an ancient record player (without sound). Even if both objects caused significant hotspots (Fig. 14), this time more people were looking to the sound object than in the test before. A reason for this could be that the gramophone was closer to the phone in the last test. Asking a participant, we got the answer that for him it was a learning process. He recognized the two phones in the test before. Therefore, this time he was looking for the object with sound.

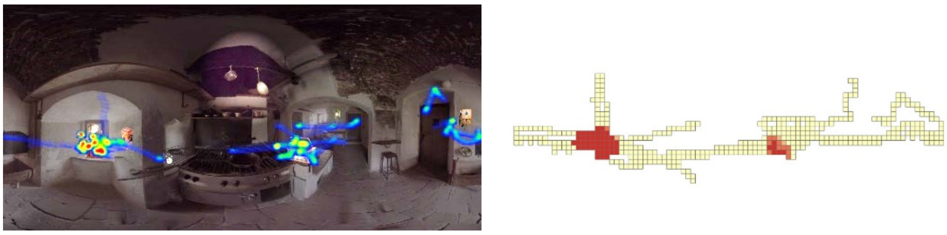

Fig. 14. Gramophone with sound (left), record player without sound (right), 2 seconds after the objects appeared

In the final step of this test section, we increased the angle to the following object – to nearly 180° (Fig. 5). A radio appeared opposite the gramophone. This time it took longer until the hotspot was built (Fig. 15 and 16).

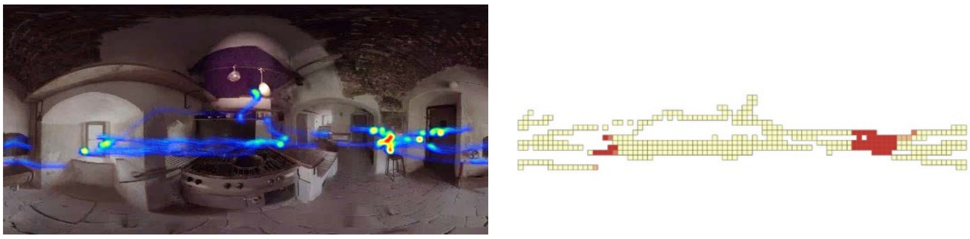

Fig. 15. Heatmap and hotspots for the radio - 2s after the radio appeared

Fig. 16. Heatmap and hotspots for the radio - 5s after the radio appeared

Scene 4

The **fourth scene** includes **lo**comotive objects **(lo)** with and without additional cues:

- a feather (lo)
- a moving cone of light (lo,l)
- a humming bee (lo,s).

When an object reaches a target position a music instrument appears and starts playing. The locomotive object disappears. The music instruments stay present and play until the end of the sequence. In this way, the scene is becoming more and more complex – visually and aurally.

Fig. 17. Moving feather. The feather hovers from the starting point (SP) to the end point (EP) - it takes 20s

The sequence starts with a moving feather (Fig. 17). Most participants followed this cue, even the time interval was relatively long (20s). In Fig. 17 the distance from the start point (SP) to the end point (EP) is illustrated. The heatmap and the significance of the hotspots can be seen in Fig. 18.

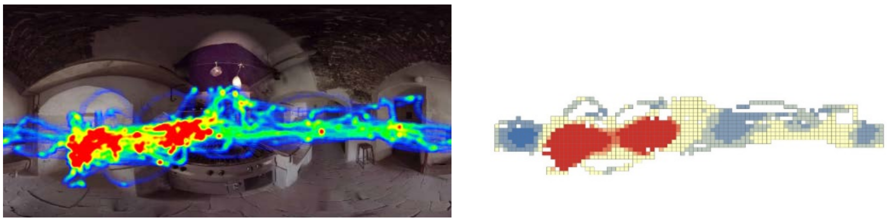

Fig. 18. Heatmap and hotspots for the moving feather

For the next test, we chose a bigger distance and time interval (47s). A light cone moves to the far side of the scene (Fig. 19).

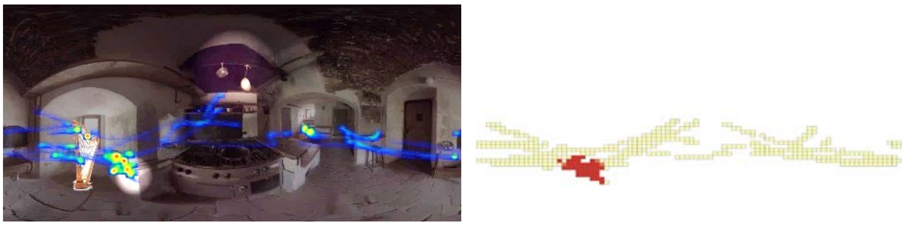

Fig. 19. The moving light cone – one hotspot at the beginning of the test interval

At the beginning of the test interval there was only one hotspot (Fig. 19). Later (around 10s) the participants were looking around, mostly to the playing instruments. Three hotspots were formed (Fig. 20). Nevertheless, mostly they returned back to the light cone and at the end of the movement there was only one hotspot again (Fig. 21).

Fig. 20. The moving light cone – three split hotspots in the middle of the test interval

Fig. 21. The moving light cone – hotspot at the end of the test interval

The last test – a moving bee with sound - was shorter (10s). Thus, the covered distance was shorter, too. As a consequence, it was more difficult to identify if the participants really followed the bee or were looking at the instruments, because the bee was flying from one instrument to the other one. Even if a hotspot was built, there is not only the bee in this area but also the instruments (Fig. 22).

Fig. 22. A hotspot around the moving humming bee

Accordingly, for all three methods investigated in scene 4, significant hotspots were built around the cues.

5 Discussion

The presented results show that diegetic cues are useful for guiding the attention of viewers in Cinematic VR.

In our first experiment (scene 1) applying the single cues simultaneously, we could not find a type of cue more efficient as the other ones. The hotspots are distributed in the room ((Fig. 6). A reason for that could be that the cues are subtle and not so flashy as in the next scenes. Furthermore, at the beginning of the movie all objects are new and cues have less effect.

Comparing combination of two cues in our second experiment (scene 2), the combinations including sounds (s,m) and (l,s) resulted in hotspots, but the lamp (m,l) was not noticed by the viewers. It needs further experiments if this result was influenced by the position of the object, which was higher than the others. In our study the combination of movement and sound had the most power for drawing the attention.

Analysing the data of scene 3, we could find, that objects connected with sound can attract the attention of the viewer, even if they are not in the FoV in the moment the sound starts. This method, is more effective if the sound is coming from the direction of the object. However, it works also in other cases. Thus, the sound itself seems to be more important than the direction of the sound.

In addition, moving objects can influence the viewing direction (scene 4). Even if the movements are for a longer period of time, viewers looking around in the meantime but mostly return to the movements. In our experiment the moving objects and lights guided the attention even without additional sound.

In **summary**, we found the following results:

- Objects connected with sounds attract more attention than without sound
- Objects connected with sounds can guide the viewing direction even if the sound is not spatial or is coming from another direction
- Moving objects or lights can guide the viewing direction even without any sounds
- It is difficult to guide the viewer at the beginning of a new scene

- Non-moving lights had no effects in our tests.

For finding these results we used heatmaps and spatial statistical methods for determining significant hotspots. In our interviews we experienced, that some people are afraid of missing something. Therefore, guiding can be helpful for making the enjoyment of Cinematic Virtual Reality more enjoyable.

Limitations

In our research, we tracked the head movements – not the eyes. Through that, we could follow the viewing direction in general, however not in detail. This was sufficient for a first approach and the most experiments led to evaluable data. Tracking the eyes instead the head would give more detailed information.

In the first scene of our tests the cues did not generate more attention than other objects in the room. We are not sure, if the reason is the weakness of the cues or the novelty of the room. In further tests we will replicate the first scene at the end of the movie.

In the last part of our experiments – we investigated if the viewer follows the bee for a short distance - we can see only a hotspot in the middle of its flight. The bee was flying between two music instruments which were visible at this time and so most people fixed the direction of the head between the instruments. The flight of the bee was short and the viewers could follow it by moving their eyes and not their heads – so the instruments stayed in the FoV. In this case eye tracking would be helpful for the analysis.

We tried to give the objects equivalent properties. However, sometimes there was an impact which we did not expect. It seems that objects in Cinematic Virtual Reality which are a little bit higher are not equivalent to objects straight ahead or a little bit lower. This might be due to the fact that we rarely look upwards in real life.

Even if we randomized the direction of view between the sections, within the sections the objects were correlated. In our experiment, we tested the sequence of cues. In the analysis of the data we found plausible additional assumptions which should be verified in further experiments where the cues are independent:

- Objects with spatial sound from the side of the viewer attract attention more and faster than objects which lie behind the viewer
- Higher objects attract attention less than objects ahead or below.

Conclusion

This research was our first step in investigating how a viewer can be guided in Cinematic Virtual Reality. We showed that sound draws the attention of the viewer. However, also objects without any sound can be used for guiding the viewer to other directions. The participants followed a locomotive cone of light and also locomotive objects.

These results can be used for integrating diegetic cues in a movie for guiding the viewer to things which are important for the story. The investigated methods require the integration of cues in the movie. This is not always possible. The viewer should have several possibilities to find an own way in the story. In our future research, we plan to examine non-diegetic methods for viewer guiding which should not decrease the presence.

In our research we looked for a method to analyze the collected data. Our data - the tracked FoV - are points on a sphere changing over time. Our approach using spatial statistical methods has proved of value and we will develop this method further in our research.

References

1. Boonsuk, Wutthigrai. *Evaluation of Desktop Interface Displays for 360-Degree Video*. Iowa State University, ProQuest Dissertations Publishing, Ames, Iowa, 2011.
2. Overschmidt, Gordian and Schröder, Ute B. (Eds.). *Fullspace-Projektion: Mit dem 360° Lab zum Holodeck*. Springer, 2013.
3. Chamier-Waite, Clea von. The Cine-poetics of Fulldome Cinema. *Animation Practice, Process & Production*, 3 (2013), 219-233.
4. Slater, Mel, Usoh, Martin, and Steed, Anthony. Depth of Presence in Virtual Environments. *Presence: Teleoperators & Virtual Environments*, 3 (1994), pp. 130-144.
5. Slater, Mel and Wilbur, Sylvia. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 6 (1997), 603-616.
6. Hendrix, Claudia and Barfield, Woodrow. The Sense of Presence within Auditory Virtual Environments. *Presence: Teleoperators & Virtual Environments*, 5 (1996), pp. 290-301.
7. Cummings, James J. and Bailenson, Jeremy N. How immersive is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, 19 (2016), pp. 272-309.
8. Poeschl, Sandra, Wall, Konstantin, and Doering, Nicola. Integration of Spatial Sound in Immersive Virtual Environments- an Experimental Study on Effects of Spatial Sound on Presence. In *Proceedings of the IEEE Virtual Reality Conference 2013* (Piscataway, NJ 2013), IEEE Computer Society Press, 129-130.
9. Serafin, G. and Serafin, S. Sound Design to Enhance Presence in Photorealistic Virtual Reality. In *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display* (Sydney, Australia 2004), International Community for Auditory Display.
10. Syrett, Hannah, Calvi, Licia, and van Gisbergen, Marnix. The Oculus Rift Film Experience: A Case Study on Understanding Films in a Head Mounted Display. In *Intelligent Technologies for Interactive Entertainment: 8th International Conference* (Cham 2016), Springer, 197-208.

11. Goldstein, E. B. *Sensation and Perception. 8th International Edition*. J. Hague & J. Perkins (Eds.), Belmont, California: Wadsworth, 2010.
12. Coren, S., Ward, L. M., and Enns, J. T. *Sensation and Perception. 5th*. Harcourt Brace College Publishers, San Diego, 1999.
13. Veas, Eduardo E., Mendez, Erick, Feiner, Steven K., and Schmalstieg, Dieter. Directing Attention and Influencing Memory with Visual Saliency Modulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)* (Vancouver,Canada 2011), ACM Press, 1471-1480.
14. Van der Burg, Erik, Olivers, Christian N. L., Bronkhorst, Adelbert W., and Theeuwes, Jan. Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*, 34 (2008), 1053-1065.
15. Hoeg, Emil R., Gerry, Lynda J., Thomsen, Lui, Nilsson, Niels C., and Serafin, Stefania. Binaural sound reduces reaction time in a virtual reality search task. In *IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)* (Los Angeles, CA 2017), IEEE, 1-4.
16. Lin, Yen-Chen, Chang, Yung-Ju, Hu, Hou-Ning, Cheng, Hsien-Tzu, Huang, Chi-Wen, and Sun, Min. Tell Me Where to Look: Investigating Ways for Assisting Focus in 360° Video. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado 2017), ACM, 2535-2545.
17. Brown, C., Bhutra, G., Suhail, M., Xu, Qi., and Ragan, E. D. Coordinating Attention and Cooperation in Multi-User Virtual Reality Narratives. In *2017 IEEE Virtual Reality* (Los Angeles, CA 2017), IEEE, 377-378.
18. Pausch, R., Snoddy, J., Taylor, R., Watson, S., and Haseltine, E. Disney's Aladdin: First Steps toward Storytelling in Virtual Reality. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques* (New Orleans, LA 1996), ACM, 193-203.
19. Sheikh, A., Brown, A., Watson, Z., and Evans, M. Directing Attention in 360-Degree Video. In *IBC 2016 Conference* (Amsterdam 2016), IET, 29-38.
20. Nielsen, L. T., Møller, M. B., Hartmeyer, S. D., Ljung, T., Nilsson, N. C., Nordahl, R., and Serafin, S. Missing the Point: An Exploration of How to Guide Users' Attention during Cinematic Virtual Reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich,Germany 2016), ACM, 229-232.
21. Getis, Arthur and Ord, J. Keith. The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24 (1992), pp. 189-206.