Exploring the Optimal Point of View in Third Person Out-of-Body Experiences

Thomas Kosch[1], Robin Boldt[2], Matthias Hoppe[2], Pascal Knierim[1], Markus Funk[1]

University of Stuttgart (Pfaffenwaldring 5a, 70569 Stuttgart, Germany)
firstname.lastname@vis.uni-stuttgart.de[1]
{boldtrn, hoppems}@studi.informatik.uni-stuttgart.de[2]

ABSTRACT

The way human perceive their environment depends strongly on the senses they have. One of them is the natural first person view using their eyes, which is limited regarding the field of view, perceived wavelengths or angular resolution. A third person view, which is often used in video games to provide a better visibility of the avatar and its surroundings, could overcome the limitations mentioned above to change the way humans visually perceive their environment. In this paper, we use an existing prototype consisting of a head-mounted display with an external camera to investigate the impact of a third person view regarding user experience and the evaluation of viewing angles. Besides manual view positioning by a user, we used an autonomic navigation task to compare both approaches. Results show that autonomic positioning shows significantly less errors than a manual positioning of the camera.

ACM Classification Keywords

H.5.1 Information Interfaces and Presentation (e.g., HCI): Multimedia Information Systems; Artificial, augmented, and virtual realities

Author Keywords

Third Person View; Virtual Reality; Out-of-Body Experience

INTRODUCTION

A first person view (1PV) is the natural way for humans to perceive their environment visually. Since the 1PV is constitutional, visual demanding fine grained actions can be accomplished. However, the fixed viewpoint makes it hard to look around or to see certain body parts (e.g. the head). Video games overcome this disadvantage with a third person view (3PV), which provides a better vision and more control to the player through a better perception of their in-game environment [1, 8, 13, 17].

PETRA '16, June 29-July 01, 2016, Corfu Island, Greece.
©2016 ACM. ISBN 978-1-4503-4337-4/16/06...$15.00
DOI: http://dx.doi.org/10.1145/2910674.2910720

The usage of 3PV is also applicable for real-life use-cases. For example people performing shadowboxing or dancing are using mirrors for analysis purposes to increase their overall performance. A 3PV allows dynamic viewing angles while performing these actions. Another use-case are medical applications. For example, paraplegic which are not able to move can be supported by a 3PV. People suffering from vertebral fractures which are influenced in their head movements can use a 3PV to look around. Visual support using 3PV can be also achieved when using machines (e.g. operating a wheelchair or parking a car). However, using a 3PV as known from video games in daily life is not common yet, although it can be built easily with available technologies using a camera and a head-mounted display (HMD). The camera can be carried at any position behind or in front of a person's body and the video stream of the camera can be redirected to the HMD. This approach becomes scalable when using low-cost VR headsets, that turn a regular smartphone into an HMD (e.g. Google Cardboard[1]). However, there are a number of possible viewpoints that a camera can provide. Furthermore, the position is not limited to a stick-worn camera, as an Unmanned Aerial Vehicle (UAV) is also able to carry a camera.

In this paper, we evaluate the position of the camera that users find most suitable for navigating through the real world using a 3PV. The contribution of this paper is twofold: (1) Through a lab study, we compare automatic positioning against manual positioning of a camera viewpoint in a 3PV scenario and (2) we evaluate the favored viewpoint which is estimated by the participants.

RELATED WORK

Previous research has proposed several systems that diversify the human's visual perception. In general, previous work can be divided into two categories. One approach is to (1) extend the user's field of view (FoV); the other is to (2) control the angle and position of the view.

Extending the FoV to increase the range of visual perception has been addressed by various researchers before. Ardouin et al. [2] developed a 360° panorama camera which allows a user to stream a video in real-time onto an HMD. The video is streamed directly to the HMD, allowing the wearer to see

[1] www.google.com/get/cardboard, (last access 04-03-2016)

Figure 1. The WOz approach (A1) allows the participant to see their full body (A2). The stick approach (B1) enables a better control of view (B2).

the complete environment in real time. Omnidirectional images were captured using catadioptric sensors. Fan et al. [6] developed a prototype, which uses a front and back camera, showing the natural FoV. The system switches to the back view if movement in the back is detected. The video is streamed to an HMD in real-time. Another approach to extend the FoV was done by Orlosky et al. [11] using wide angle lenses. They extended the FoV from 180° panorama to 238° panorama. Further, they conclude that extending the human's FoV is possible, but with a lack of details. A non-visual extension of the user's FoV was proposed by Mateevitsi et al. [10]. In their Spider Sense project, they use vibrotactile feedback as sensor substitution for the vision. Possible obstacles are felt as vibration while walking. Creem-Regehr et al. analyzed the influence of human behavior when changing the FoV using an HMD [5, 9]. Changing the view position influences the participants' distance judgment. This could lead to a falsified distance judgment when using a wide-angle lens on the camera.

Changing the camera position to modify the viewpoint has been done before. Bowman et al. [4] investigated the motion of humans in immersive virtual environments from different viewpoints. They built a framework, which was displayed on an HMD, and tested different movement taxonomies. Changing the viewpoint rapidly disoriented participants, who needed some time to regain orientation. Raising the number of objects in motion increased the user's cognitive load. On the other hand, Rapp and Gena [12] compared 3PV to 1PV, to observe the benefits of both approaches. Their experiment comprised several tasks, using 3PV and 1PV in a virtual environment. Their findings show that searching and finding objects is easier in 1PV. 3PV is suited better for orientation. Furthermore, Salamin et al. [14, 15] built an experimental prototype, consisting of a backpack in combination with a stick and a camera, fixated at one end of the stick. The stick was mounted on the backpack, which allows the user to use both hands. The system was ready to use after the correct viewing angle for the wearer was set. Although the system provided a 3PV, it was not possible to adjust the viewing angle during runtime, since the stick had a static behavior. An extended version of the system used multiple cameras mounted in the environment and one wearable camera on a stick [16]. The view switched from the stick-camera to the environment-camera if the user entered a room that contains an environment-camera.

Overall, related work focused on extending the human's vision by either increasing the field of view or experimenting with alternative camera positions. However, previous work did not investigate different dynamic 3PV camera positions during runtime. Therefore, we focus on dynamic camera positioning in 3PV.

SYSTEM

We used an existing prototype consisting of a camera stick and an HMD (see Figure 1) [3]. The stick has a Creative Socialize HD webcam with a resolution of 640×480 mounted at one end of it. Additionally, a cord was connected to the webcam to enable users during runtime to adjust the vertical viewing angle of the camera. Since the human horizontal FoV is close to 180° [7], we attached a wide-angle lens to the webcam to increase the FoV to 180°. We combined a Samsung Galaxy S4[2] and a Google Cardboard as HMD. A Lenovo T440s was used to connect the HMD and the webcam. The notebook was used to retrieve the video from the webcam and to stream it to the HMD afterwards. The video stream itself is sent to the HMD using VR Streamer[3].

EVALUATION

We conducted a user study to find the best-perceived viewpoint in 3PV. We compared two different viewpoints and camera positioning approaches. Either the participant was allowed to manipulate the camera position on its own using the stick or the camera position was adjusted by the experimenter, the so called Wizard of Oz (WOz). In a first part of an experiment, participants had to walk through a maze using either the stick or WOz. In succession, participants were asked to walk a predefined path around our university's building. In a second part of the experiment, participants were asked to try out different camera positions, estimate distances, and handle obstacles of everyday life.

---

[2] Running Android 4.4.2

[3] www.swatterco.com/vr_streamer.php (Version 1.2, last access 04-03-2016)

Method

The experiment was conducted using a repeated measures design with two levels. The independent variable was the camera positioning approach: Either the camera was positioned by the user using the camera stick (see Figure 1, B1) or the camera was positioned by the experimenter (see Figure 1, A1). The conditions were counterbalanced to eliminate learning effects. As dependent variables, we measured the task completion time (TCT) and the errors made by the participant when walking through the maze. We counted an error when the participant overstepped the borders of the maze.

Apparatus

The task of the lab study was to walk through a maze. The length of the maze was $30\,m$. The width of the path in the maze was about 0.8 m. Furthermore, we used boards as physical borders and applied tape on the floor to indicate the path of the maze.

Procedure

After explaining the course of the study, we introduced the participants to our prototype. We let them walk through a test corridor to get used to viewing themselves in 3PV. When the participant felt comfortable using the system, they were asked to complete the maze using either automatic camera positioning or manual camera positioning. The starting and end position was the same for each condition. TCT and errors were measured as soon as the participant started walking from the starting point. The measurements were completed after the participant reached the endpoint. We repeated the procedure for both conditions. To further experiment with the viewpoint, we asked the participants to walk a pathway around our university's building. The pathway is 400 m long. The participants were asked experiment with different camera angles and camera positions to find the optimal viewpoint for everyday tasks. Afterwards, we collected qualitative feedback through semi-structured interviews. At the end, we collected subjective feedback about the preferred viewing positions of the camera, using a Likert scale ranging from 1 (very bad) to 5 (very good). The study took about 30 minutes per participant.

Participants

We recruited 20 participants (10 female) aged between 18 and 43 years ($M = 22.95, SD = 5.28$) through our university's mailing lists. A total of 9 participants were wearing glasses.

Results

Results of the TCT indicate that on average using the stick (M = 91.35, SD = 34.80) was faster compared to WOz (M = 100.30, SD = 41.34). A Shapiro-Wilk test showed that both samples are not normally distributed (stick: $p = .03$, WOz: $p < .01$). A Kruskal-Wallis rank sum test ($\chi^2(1,40) = .4396, p = .51$) could not show a significant difference between the conditions. The average number of errors using the stick was 2.95 ($SD = 2.04$) while 1.85 ($SD = 2.48$) errors were made using the WOz approach. A Shapiro-Wilk test showed that the errors made using WOz are not normally distributed ($p < .01$). In contrast, using the stick, the error data are normally distributed ($p = .11$). A Kruskal-Wallis rank sum test revealed a significant difference between the samples

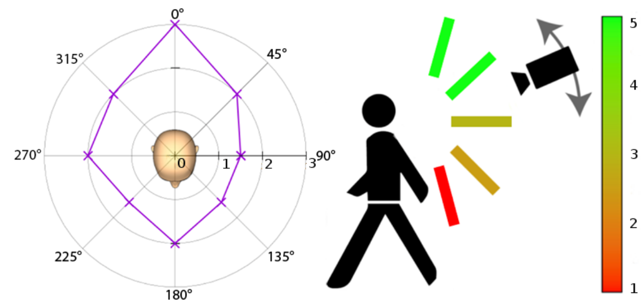

Figure 2. Left side: Preferred horizontal camera positions. Right side: Preferred vertical camera positions.

($\chi^2(1,40) = 4.25, p = .04$). Considering the two approaches, 50% of the participants preferred the WOz approach and 20% of the participants preferred manually changing the viewpoint using the stick. All, except two participants, are sure that with more experience using a third person VR system they will be more efficient and comfortable.

The Likert questionnaire contained eight different horizontal camera positions (see Figure 2): From behind the head in an angle of 0° ($M = 2.7, SD = .98$), behind the head from the left in an angle of 45° ($M = 2.35, SD = .59$), from left of the head in an angle of 90° ($M = 1.7, SD = .86$), in front of the head from the left in an angle of 135° ($M = 1.7, SD = .86$), in front of the head in an angle of 180° ($M = 1.75, SD = .79$), in front of the head from the right in an angle of 225° ($M = 1.6, SD = .75$) from right of the head in an angle of 270° ($M = 1.9, SD = .91$), and behind the head from the right in an angle of 315° ($M = 2.15, SD = .88$). A Shapiro-Wilk test showed that all 8 measures are not normally distributed (0°:$p = .04$; 45°- 315°:$p < .01$). A Kruskal-Wallis rank sum test revealed a significant difference between the samples ($\chi^2(7,160) = 27.67, p < .01$). The Likert questionnaire contained five different vertical camera positions (see Figure 2): from above in an angle of 180° ($M = 2.4, SD = .96$), from above and behind the head in an angle of 135° ($M = 2.18, SD = .78$), from behind the head in an angle of 90° ($M = 1.88, SD = .89$), from behind and below the head in an angle of 45° ($M = 1.43, SD = .49$), and from behind and below in an angle of almost 0° ($M = 1.3, SD = .57$). A Shapiro-Wilk test showed that 4 measures are not normally distributed (180°: $p = .01$; 135°: $p = .09$; 90°: $p < .01$; 45°:$p < .01$; 0°: $p < .01$). A Kruskal-Wallis rank sum test revealed a significant difference between the samples ($\chi^2(4,100) = 24.4, p < .01$).

Discussion

We conducted a study that compared two ways of positioning the camera in space to navigate through a maze. While we cannot find a significant effect of the camera positions on the TCT, we found a significant effect on the errors. 17 participants stated that seeing their feet while navigating is a major benefit of WOz. When using the stick, some participants positioned the stick somewhere over the head to see their feet. Others put the camera as far behind their head as possible to increase the FoV. When encountering obstacles or when changing the direction most participants switched to the overhead camera position. The result of the Likert questionnaire supports the qualitative statements of the participants that most like to see their feet. In general, a camera position that shows the feet but still provides a lot of overview was favored.

Limitations

In the WOz approach the camera was moved by the experimenter. The adjustments made by the experimenter were delayed, since the participants sometimes moved odd and the experimenter needed a short time to fix the viewing angle for the participants. Therefore, we believe that the task completion time could be improved for the WOz approach when automatically adjusting the camera position. Another limitation is the received camera image while performing the WOz approach which depends highly on the position the experimenter is holding the camera. The position of the camera hold by the experimenter differs for every participant, which may influence the user experience and the amount of errors made during the user study. The problem can be solved by using an UAV, which is able to set up the optimal viewpoint automatically.

CONCLUSION

In this paper, we used a low-cost system to provide a third person view for real life settings. In a lab study, we compared an automatic camera positioning using a WOz towards a manual positioning of the camera. The results indicate when using an automatic positioning of the camera users made significantly less errors than using a camera-mounted stick. In a qualitative evaluation, we found that top right, behind the user's head is the preferred viewpoint of a camera providing a real life third person experience. In future work, we want to create a fully automatic positioning using UAV's to orient a camera automatically depending on the users movement. Additionally, we want to create an experiment where the camera has a fixed position during a navigation task. This enables more precise measurements regarding TCT, errors made and comfort of participants for different camera viewpoints.

ACKNOWLEDGMENTS

This work is funded by the German Research Foundation within the SimTech Cluster of Excellence (EXC 310/2).

REFERENCES

1. Ernest Adams. 2014. *Fundamentals of game design*. Pearson Education.
2. Jérôme Ardouin, Anatole Lécuyer, Maud Marchal, Clément Riant, and Eric Marchand. 2012. FlyVIZ: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proc. VRST*. 41–44.
3. Robin Boldt, Matthias Hoppe, Thomas Kosch, Markus Funk, Pascal Knierim, Bastian Pfleging, and Niels Henze. 2015. Towards an Optimal Viewpoint in Third-Person out-of-body Experiences. *MUC'15* (2015).
4. Doug Bowman, David Koller, Larry F Hodges, and others. 1997. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *IEEE VR'97*. 45–52.
5. Sarah H Creem-Regehr, Peter Willemsen, Amy A Gooch, and William B Thompson. 2005. The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception* 34, 2 (2005), 191–204.
6. Kevin Fan, Jochen Huber, Suranga Nanayakkara, and Masahiko Inami. 2014. SpiderVision: extending the human field of view for augmented awareness. In *AH'14*. 49.
7. E Bruce Goldstein. 2009. *Encyclopedia of perception*. Sage Publications.
8. François Dominic Laramée. 2002. *Game design perspectives*. Charles River Media, Inc.
9. Markus Leyrer, Sally A Linkenauger, Heinrich H Bülthoff, Uwe Kloos, and Betty Mohler. 2011. The influence of eye height and avatars on egocentric distance estimates in immersive virtual environments. In *Proc. SIGGRAPH*. 67–74.
10. Victor Mateevitsi, Brad Haggadone, Jason Leigh, Brian Kunzer, and Robert V Kenyon. 2013. Sensing the environment through SpiderSense. In *Proc. AH'13*. 51–57.
11. Jason Orlosky, Qifan Wu, Kiyoshi Kiyokawa, Haruo Takemura, and Christian Nitschke. 2014. Fisheye vision: peripheral spatial compression for improved field of view in head mounted displays. In *Proc. SUI'14*. 54–61.
12. Amon Rapp and Cristina Gena. 2014. Immersion and involvement in a 3D training environment: Experimenting different points of view. In *Proc. CIVEMSA'14*. 18–23.
13. Richard Rouse III. 1999. What's your perspective? *ACM SIGGRAPH Computer Graphics* 33, 3 (1999), 9–12.
14. Patrick Salamin, Daniel Thalmann, and Frédéric Vexo. 2006. The benefits of third-person perspective in virtual and augmented reality?. In *Proc. VRST'06*. 27–30.
15. Patrick Salamin, Daniel Thalmann, and Frédéric Vexo. 2008. Improved Third-Person Perspective: a solution reducing occlusion of the 3PP?. In *Proc. VRCAI'08*. 30.
16. Patrick Salamin, Daniel Thalmann, and Frédéric Vexo. 2010. Providing the best third-person perspective to a video-through HMD. In *Proc. ICCAE'10*.
17. Ellen L Schuurink and Alexander Toet. 2010. Effects of third person perspective on affective appraisal and engagement: Findings from SECOND LIFE. *Simulation & Gaming* 41, 5 (2010), 724–742.