Influence of Narrative Elements on User Behaviour in Photorealistic Social VR

Silvia Rossi
UCL
London, UK
s.rossi@ucl.ac.uk

Irene Viola
CWI
Amsterdam, the Netherlands
irene@cwi.nl

Jack Jansen
CWI
Amsterdam, the Netherlands
jack.jansen@cwi.nl

Shishir Subramanyam
CWI
Amsterdam, the Netherlands
s.subramanyam@cwi.nl

Laura Toni
UCL
London, UK
l.toni@ucl.ac.uk

Pablo Cesar
CWI, TU Delft
Amsterdam, the Netherlands
Delft, the Netherlands
p.s.cesar@cwi.nl

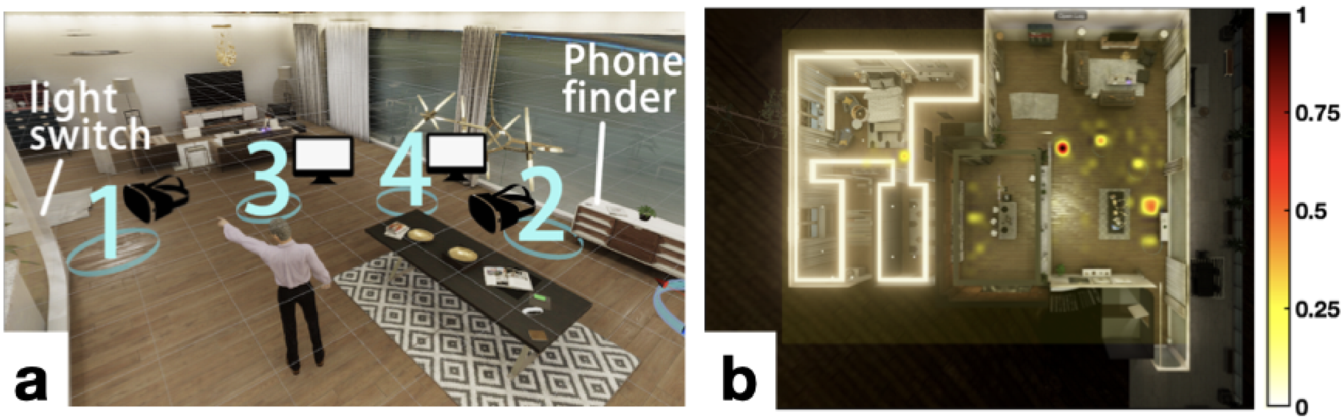


Figure 1: a) Living room with indicated the starting position of each user: 1 and 2 are HMD users while 3 and 4 are desktop users. There are also two interactive objects (*i.e.*, the light switch on the left and phone finder on the right) and the main virtual character, detective Sarge. b) Floor map of the virtual house with the user heatmap of main locations visited over time.

ABSTRACT

Social Virtual Reality (VR) applications represent a big step forward in the field of remote communication. Social VR provides the possibility for participants to explore and interact with virtual environments and objects, feelings of a full sense of immersion, and being together. Understanding how user behaviour is influenced by the shared virtual space and its elements becomes the key to design and optimize novel immersive experiences. This paper presents a behavioural analysis of user navigating in 6 degrees of freedom social VR movie. Specifically, we analyse 48 user trajectories from a photorealistic telepresence experiment, in which subjects watch a crime movie together in VR. We investigate how users are affected by salient agents (i.e., virtual characters) and by narrative elements of the VR movie (i.e., dialogues versus interactive part). We complete our assessment by conducting a statistical analysis of the collected data. Results indicate that user behaviour is affected by different narrative and interactive elements. We conclude by presenting our observations and drawing conclusions on future paths for social VR experiences.

This work has been supported by Royal Society under grant IES R1180128 and by Cisco under Cisco Research Center Donation scheme.

MMVE'21, Sept. 28-Oct. 1, 2021, Istanbul, Turkey

ACM ISBN 978-1-4503-8436-0/21/09…$15.00
https://doi.org/10.1145/3458307.3463371

CCS CONCEPTS

• **Information systems** → **Multimedia streaming**; • **Human-centered computing** → **User studies**; **Virtual reality**; Collaborative interaction.

KEYWORDS

User Analysis, 6-DOF, Social VR, Point Cloud, Immersive Movie

ACM Reference Format:
Silvia Rossi, Irene Viola, Jack Jansen, Shishir Subramanyam, Laura Toni, and Pablo Cesar. 2021. Influence of Narrative Elements on User Behaviour in Photorealistic Social VR. In *International Workshop on Immersive Mixed and Virtual Environment Systems (MMVE'21), September 28, 2021, Istanbul, Turkey*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3458307.3463371

1 INTRODUCTION

Virtual Reality (VR) applications are going through a rapid evolution of technology, getting integrated in daily-life devices such as smartphones and laptops. Therefore, it is possible to imagine that in the near future video calls will give users a completely different experience than now: people will be able to chat, walk together, interact with virtual objects and watch events such as concerts or movies together in a common virtual environment [15]. This is what *social VR* applications are pledging to enable, becoming a promising tool of the near future remote communications [13]. VR applications overpassed the traditional paradigm of passively consuming multimedia content, enabling a sense of immersiveness and interaction by placing the users at the centre of the action; a further step toward fully immersive services has been attempted by social VR. This emerging remote communication tool is indeed aiming at overstepping current remote communications through 2D screens, enabling instead virtual co-presence of more users within the same virtual environment and allowing body interactions similarly to face-to-face communication [7, 13, 14]. The new challenge is therefore to increase the realism in the virtual experience and interaction. Therefore, the key aspect that needs to be fully understood in order to advance in this technology is the user, the main director of the virtual experience.

Emerging Social VR platform, such as *Facebook Horizon*[1] and *Mozilla Hubs*[2], are rapidly growing in popularity. For instance, the latter one has been used in many recent academic events (*i.e.*, conferences: IEEE VR 2020[3], ACM IMX 2020[4], QoMEX 2020[5] and ACM CHI 2020 Social VR workshop [13]). In most of these applications, participants interact among each other by taking part of 3D virtual spaces through a computer-generated and customised avatar. Physical displacement and proxemic interactions in virtual environments have been analysed to investigate which social cues are the most influencing and therefore are needed to ensure presence and immersion [12, 27]. Moreover, many works have investigated the advantages to have a more realistic self-representation versus a 3D avatar providing a higher degree of immersion and presence in the VR experience [8, 11]. Therefore recently, a more natural self-representation of participants has been introduced, thanks to real-time acquisitions and reconstructions of point clouds by depth cameras [9]. Not to be neglected is the technological aspect of these new applications. Social interactions and photorealistic representations come with both computational and bandwidth overhead for transmission and rendering [21]. Analysing user behaviour in terms of viewing angle and head trajectories, among other objective measurements, represents the first step towards the building of real-time and realistic VR systems at large that can optimize delivery based on user-centred adaptation [24, 26].

In this paper we focus on a better understanding of how people interact with a virtual environment and other users within it. In details, we present a first attempt of objective behavioural analysis in a 6-Degrees-of-Freedom (DoF) social and interactive VR movie. We based our analysis on navigation trajectories collected in a novel type of interactivity crime movie: four users, either equipped by head-mounted display (HMD) or desktop computer and a controller, were watching together a VR crime solving movie [17]. Figure 1 (a) show a snapshot of the living room in which the story mainly takes place, with its main virtual character. Using hyper-realistic self-user representations and enabling activities such as handling virtual objects and talking with characters, the main novelty of this experience is to provide a unique sense of immersion and social connectedness going beyond to a collaborative game experience [17]. We selected this content because highly representative of social VR content, in which users are free to move in the 3D space (6-DoF), but also to interact among themselves within a guided content. These features however lies key challenges that we aim at overcoming in this work. As first we need to consider both the new physical settings and locomotion functionalities given to users in a 6-DoF system. The user now not only can select the portion to be displayed by rotating the head as in a 3-DoF system but can also move inside the virtual environment changing the distance and perspective with the displayed content. The second challenge to consider is the guided and interactive behaviour that is the new social and interactivity features of the application that brought an added level of dynamics. For instance, during the experience participants are asked to make simple tasks, such as to look for and to press a button; their action influences directly the narration of the story since the time to solve the task is not fixed. Therefore, we compare user behaviour in terms of spatial displacements in the virtual environment and viewing direction with respect to movie characters and other participant within the whole experience. In particular, we show how much narrative elements of the movie, such as virtual characters movements or request of interactions, influence user behaviour.

In the following, we first present related work focusing on behavioural analysis in a VR system. More details about the crime movie and experimental setup are given in Section 3. Section 4 presents the core of our analysis and discussions on our results highlighting key aspects that influence the user's behaviour in the social VR experience. Finally, a summary of our work and findings is provided in Section 5.

2 RELATED WORK

Depending on the locomotion functionalities and the type of content representation, navigation in immersive media environments can range from 3-DoF to 6-DoF. In the first case, users are placed at the center of a spherical content, and by changing their viewing angle, they are able to select the portion of the content they want to visualize at any given moment. In case of 6-DoF, users are additionally able to change their position in the 3D environment, which is now populated with volumetric objects that can be observed from any position and viewing angle. The 6-DoF scenario promises a much more natural interaction and exploration of the 3D scene than its 6-DoF counterpart, thus increasing the feeling of presence and immersiveness [2] and opening the gate to VR innovations. Incorporating information about user behaviour in user-centric systems is of paramount importance to optimize transmission and rendering [19, 26], highlighting the need to understand user behaviour in 3- and 6-DoF VR settings. In the first scenario, users trajectories

[1]https://www.oculus.com/facebook-horizon/
[2]https://hubs.mozilla.com
[3]https://ieeevr.org/2020/
[4]https://imx.acm.org/2020/
[5]https://www.qomex2020.ie

has been intensely analysed by different tools: angular velocity, frequency of fixation, and mean exploration angles, saliency maps [4, 22] but also through clustering approach and information theory metrics [16, 18, 20]. In the past, preliminary study on locomotion and display technology were presented for CAVE environments [3, 25] and more recently 6-DoF trajectories have been analysed in terms of angular velocity [1, 23]. Despite the great potential, the study of user behaviour has been overlooked focusing only on a general characterisation and none specific tool or procedure has been proposed for 6-DoF trajectories.

3 A SOCIAL VR MURDER MYSTERY MOVIE

We contextualize our experiment in a social VR setting, in which 4 users are called to experience an immersive movie, occasionally being asked to take part of the story. The scenario allows for users with 6-DoF navigation within a photorealistic 3D environment, which is populated by three virtual characters. In the following, we describe the movie timeline, the setup used during experiments and finally, we show some general performance of the system.

3.1 Movie plot

The interactive and immersive VR movie used in the experiments is fragment of a murder mystery investigation [17], led by detective Sarge Hoffsteler and his assistant, Rachel Tyrell. The users (4 in total) help the investigation. The victim is Elena Armova, who lived in a luxury apartment in central London. The investigation is split into 3 chapters, as depicted in Figure 2. In Chapter 1, the 4 users are placed in the virtual living room of the victim Ms. Armova, adopting an initial fixed positions indicated in Figure 1 (a). Participants mainly listen to a rendered victim interrogation, which is possible thank to a futuristic machine based on artificial intelligence. There are also two moments in which users are asked to interactively interact with objects in the scene: *(1)* user 1 is asked to switch on the light, *(2)* user 2 has to pick up a phone finder controller and press the button. This split Chapter 1 into three narrative moments where mainly virtual characters are talking and walking around the scene (*narrative* label in Figure 2) interleaved with two moments of active tasks for the user (*task* label). At the end of Chapter 1, users are split into two groups: user 1 and 3 are conducted to the virtual kitchen with detective Sarge, whereas user 2 and 4 are led to the virtual bedroom with Rachel and Ms. Armova. In both rooms, participants are listening to narrative dialogues of virtual characters. At the end of Chapter 2, the users are brought back together in the virtual living room for the final chapter, where detective Sarge describes how the murder has been solved.

3.2 Experimental Setup

A low-latency volumetric video delivery pipeline, based on point cloud representation, was used to place each participant into the virtual scene [9]. Each user was captured using 3 Kinect Azure devices, placed in a circle around them, 120° apart from each other. This allowed each participant to be captured from multiple angles, ensuring a photorealistic representation while they interacted with the scene.

Two devices were used to visualize the social VR experience: users 1 and 2 were equipped with Oculus Rift HMDs, complete with controllers, whereas users 3 and 4 could watch the scene through 50-inch monitors, and could navigate using gaming joysticks. For HMD users, teleportation was enabled in key locations of the scene, as physical locomotion was restricted due to the acquisition setup, whereas for desktop users, movement was enable through the gaming joystick. Due to the configuration of the controllers, only HMD users were able to engage with the interactive elements in the scene.

A total of 48 participants was recruited for the experiment, resulting in 12 social VR sessions. The number of users was selected to ensure at least 24 participants per condition (HMD vs desktop). The sample size was determined using software G*Power [5], considering the between-subject design, assuming a large effect size (d = 0.7) and setting $\alpha = 0.05$ and desired power $1 - \beta = 0.75$. The participants were between 21 and 56 years old ($\mu = 34.9, \sigma = 10.3$). The gender distribution was balanced (23 males, 25 females). Users were randomly assigned to each device hence to the initial position of the experiment. All of them were fluent in English, and had no motor or visual impairment. Participants knew themselves always in advance, in detail there were always at least two groups of friends or relatives per each experiment.

Before and after the virtual experience, semi-structured interviews were also conducted to collect explicit feedback from users. Analyzing these explicit feedback is however beyond the scope of this paper, which is rather focused on building new metric to analyse users' behavior and deduce implicit feedback. Each experimental session lasted approximately 60 minutes and consisted of the following main parts: **Part 1** *(10 minutes)*: explanation of the experiment including main goal, procedure and description of the movie characters. This phase includes filling-in questionnaire too; **Part 2** *(10 minutes)*: training phase to let participants familiarise with devices and to interact with each other and virtual object within the virtual environment; **Part 3** *(10 minutes)*: virtual movie experiment; **Part 4** *(20 minutes)*: questionnaire phase to evaluate the social VR experience (*e.g.*, filling of presence/immersion, visual quality). **Part 5** *(10 minutes)*: a semi-structured group interview with all 4 participants. During the sessions, the position and rotation of the camera objects associated with each users were recorded at 30 Hz. From the logged data it was possible to compute the latency between encoding and rendering, for each device under use. The data was obtained at the granularity of one second, to avoid disrupting the performance of the system. On average, the point count for each frame was 86342 points per cloud (25 percentile: 82412.5; median: 90268; 75 percentile: 95727.5). The observed framerate for all the representations was on average 9.08 frames per second (25 percentile: 7.5; median: 8.9; 75 percentile: 10.6). The observed latency in all sessions was generally lower than 1 second, and was remarkably smaller for self-representation. In particular, mean latency for self-representation was 0.1147 seconds (25 percentile: 0.079; median: 0.105; 75 percentile: 0.136), whereas for the rest of the cases, it amounted to 0.5526 seconds (25 percentile: 0.408; median: 0.538; 75 percentile: 0.689).

4 BEHAVIOURAL ANALYSIS

In the following, we first present a general analysis of users' movements within the entire virtual environment and movie. In order to better understand how user behaviour changed over time as the

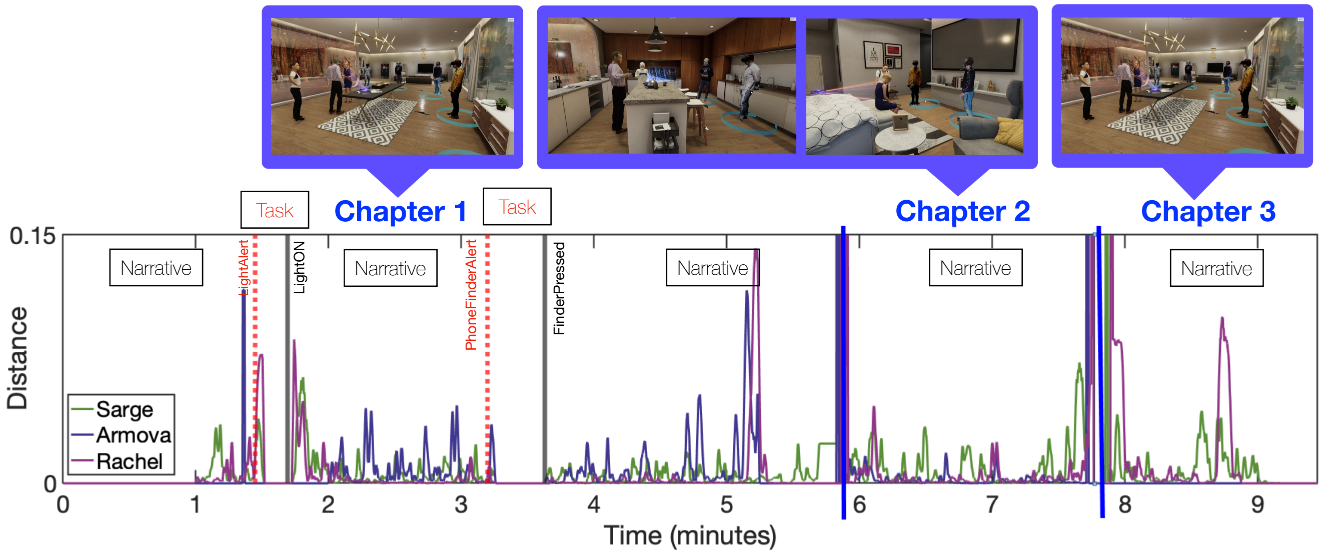


Figure 2: Timeline of the VR murder mystery movie, and spatial movements of virtual characters over time. Distance is computed with respect to the position in the previous frame. A screenshot per each chapter is also reported.

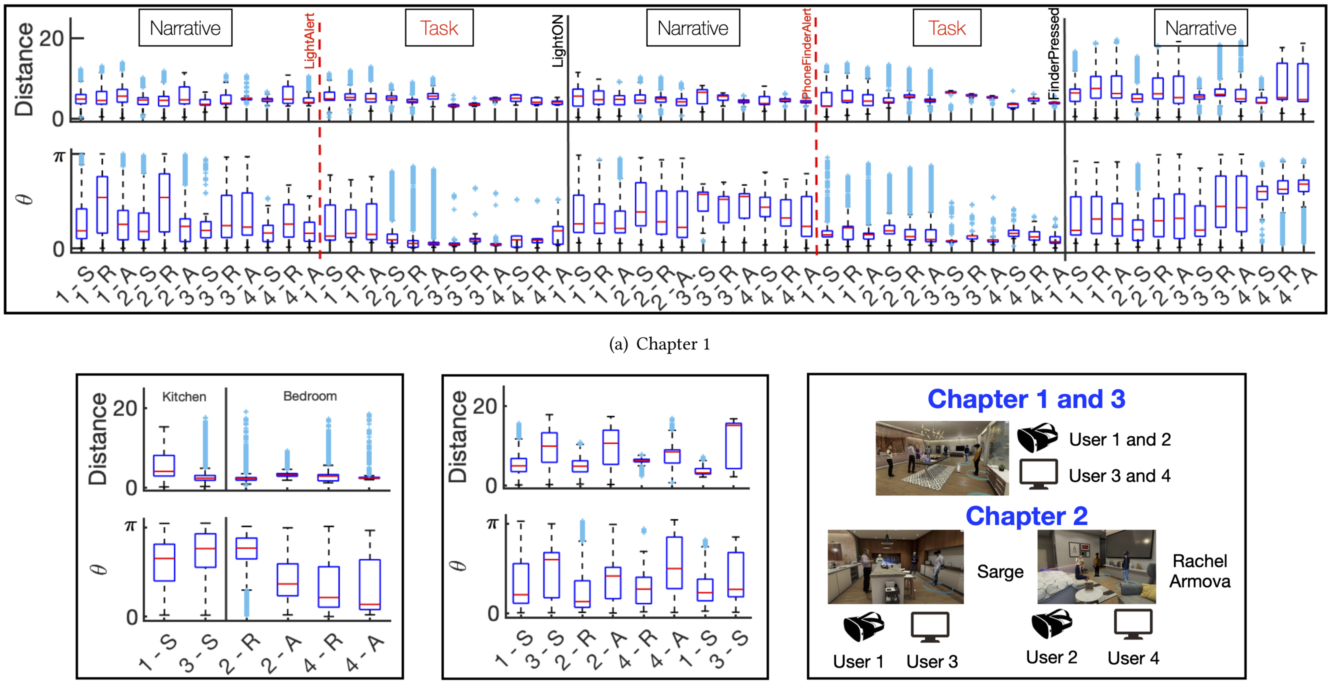
Figure 3: User motion based on both spatial and rotation movements per each session in the movie.

movie progressed, we also analyse their position and orientation with respect first to the virtual characters (avatars) and finally, to other participants.

General movements analysis. A general overview of the exploration behaviour of users is given in Figure 1 (b), which shows a heatmap of the most visited locations in the virtual house, obtained by aggregating all the position data collected in the experiment. As described in Section 3.1, large part of the movie takes place in the living room (Chapter 1 and 3), which is reflected in the figure. The most visited locations correspond to the initial positions of user 1 and 2, due to their movement restriction as HMD users (*i.e.*, only teleportation in fixed locations was allowed). More generally, the initial positions of all users are clearly visible in the heatmap; additional yellow spots outside of the predefined circles indicate most likely regions to be visited by desktop users. Whereas movement is more spread in the living room, in the other smaller room of the house, kitchen and bedroom, it appears much more spatially focused, indicating that participants were more static in these spaces. Figure 3 displays the boxplot comparison between percentage of motion exhibited by each user. For every frame, we considered the user to be "in motion" if either their relative position with respect to the previous frame changed more than 0.05 cm, or if any of their rotation angles varied by more than 0.01 rad (0.573°). Both measures were taken into account to cover both spatial exploration behavior, as well as changes in viewing angles. The percentage was then computed with respect to the total number of frames. It can be observed that desktop users (user 3 and 4) exhibit a larger percentage of motion over the course of the movie, with respect to HMD users. Motion was present in the first and last chapter with wider distributions, indicating larger variance in the way users behaved, whereas in the second chapter, users generally showed smaller variance in percentage of motion.

Movements analysis: users vs. avatars. We now compare user behaviour with respect to movements of avatars, namely Sarge, Rachel and Armova. To give an idea about avatar displacements and actions in the movie, Figure 2 shows distance covered by each character over time with respect to the previous position of the character. We can notice that during *task* phases, no movement is observed as avatars mainly waits for users to make the action. Also, a spike in the distance is observed around minute 5 and this is because Armova and Rachel leave the scene. Finally, Armova does not appear in the scene only in Chapter 1 and 2.

Equipped with this background information, we now study the distribution of relative position and orientation of the users with respect to the avatars, separately per each chapter of the movie (Figure 4). In details, the first line of each subplot depicts the distance



(a) Chapter 1

(b) Chapter 2

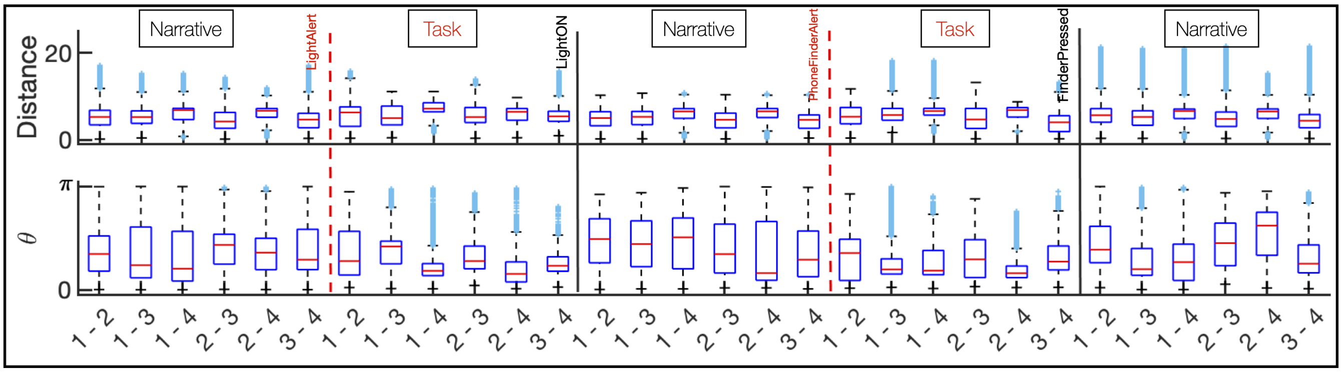
(c) Chapter 3

(d) Legend

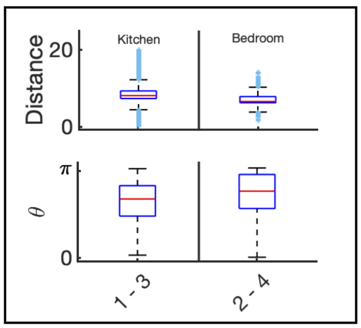
Figure 4: Distribution of spatial distances (first line of each subplot) and angles (second line of each subplot) between users and avatars per each chapter of the storytelling.

between each user and each avatar. The second line of subplots shows instead the angle between user's viewing direction and the vector which connects user and avatar at any given time. This angle indicates therefore if user is looking towards direction of character location ($\theta \rightarrow 0$) or in in the opposite direction ($\theta \rightarrow \pi$). In Chapter 1 (Figure 4 (a)), it is therefore interesting to notice the different user behaviour between *narrative* and *task* parts of the story. During the three narrative parts, users tend to further explore the environment around them: even if on average, users are looking in the direction of the three virtual characters, their variance indicates a non-uniform behaviour over time, suggesting that participants where also looking around. In the two task parts, instead, the distribution of both spatial distance and angle values is narrower. In particular, for the first task, a wider distribution in terms of viewing angle, which skews further away from the avatar, can be observed for user 1, which is the one tasked with pressing the button to turn on the light. This difference might be due to the mismatching of difficulty between the two task. Indeed, participants took on average around 13.33 seconds to switch on the light, while picking up the phone finder controller and press the button required around 25.58 seconds. To validate our intuition, we perform a non-parametric Mann-Whitney statistical test between the distance and angle to avatar recorded during the narrative parts of Chapter 1, and the task parts. A non-parametric test was selected due to the non-normality of the data distribution, according to a Kolmogorov-Smirnoff test. To avoid bias induced by the large number of samples, we performed random sampling on the data, selecting N = 200 samples across the distance vector, and repeating the procedure across 200 sampling runs. We used Fisher's method [6, 10] to combine the probabilities, obtaining that the type of task has a significant effect on the distance ($\chi^2 = 3202.5$, $p < .001$) and on the viewing angle ($\chi^2 = 4771.8$, $p < .001$). In fact, distance to avatar appears to be statistically different between all sub-parts of Chapter 1, indicating varying behavior in terms of spatial movements between as the time progressed. In terms of viewing angle, however, no discernible effect is observed on different narrative parts with respect to the viewing angle (narr.1 - narr.2: $\chi^2 = 407.5$, $p = 0.387$; narr.1 - narr.3: $\chi^2 = 407.5$, $p = 0.388$; narr.2 - narr.3: $\chi^2 = 421.5$, $p = 0.221$), whereas the two tasks exhibited significantly different viewing angle distributions ($\chi^2 = 1270.6$, $p < .001$). In Chapter 2 (Figure 4 (b)) participants are moved in different rooms, kitchen and bedroom, both of them smaller compare to the initial living room. This different ambient dimension affects indeed user behaviour: participants in general are much more static compare to the previous chapter. As last observation, we notice that users in Chapter 3 (Figure 4 (c)) behave similar to narrative moments of Chapter 1: there are exploration movements both in terms of spatial and angle values.

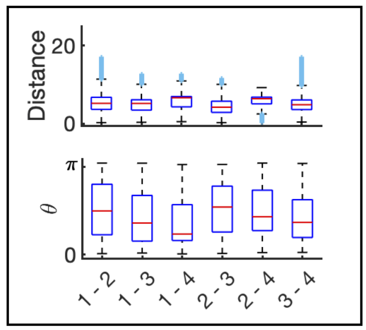
Movements analysis: users vs. users. Finally, we analyse the user's position and viewing direction with the respect to other participants. As in the previous analysis, Figure 5 depicts the spatial distance and angular difference between each couple of users per each chapter of the movie. While in the previous comparison between user and avatars, the behaviour of the latter was known and stayed constant for each experiment, in this case both users under exam have varying positions and viewing directions over



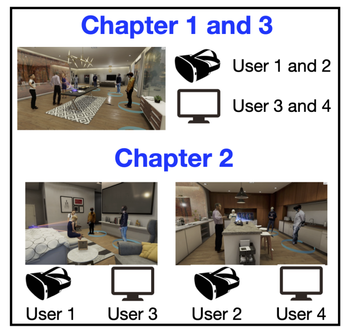
(a) Chapter 1



(b) Chapter 2



(c) Chapter 3



(d) Legend

Figure 5: Distribution of spatial distances (first line of each subplot) and angles (second line of each subplot) between couples of users per each chapter of the storytelling.

time. However, some general behaviour can be extracted also under these conditions confirming the previous findings. For instance, during task parts in Chapter 1 (Figure 5 (a)), the distribution of angle values is quite narrow for most of the couples, indicating that they moved their attention from the avatars to another participant. On the contrary, during narrative parts, these distributions are wider highlighting that users were not fixating on each other, rather exploring the scene and the virtual characters within. Statistical tests show that the effect of the task is significant on the viewing angle ($\chi^2 = 2465.6$, $p < .001$), and differences among narrative and task sub-parts of Chapter 1 are always significant ($p < .001$), with the exception of the first and the last part of Chapter 1 ($\chi^2 = 415.9$, $p = 0.282$). In terms of spatial displacements, the distance between users remains low in all the chapters of the movie. The effect of task versus narrative is significant ($\chi^2 = 1270.4$, $p < .001$), and distance always differs significantly between sub-parts ($p < .001$), with the exception of the first two narrative parts ($\chi^2 = 418.8$, $p = 0.249$).

In summary, the following observations can be deduced by the behavioural analysis carried out in this work:

- **Observation 1:** during narrative moments of the story, participants are more inclined to explore the virtual environment with general attention to virtual characters;
- **Observation 2:** the request of interactions with the content by a specific user (*e.g.*, to press a button) leads to reduced movement, while the attention is more focused on the task or on other participants;
- **Observation 3:** the size of the virtual environment in which is located the experience also affect the user's behaviour. In particular, large rooms seem to be more conductive of exploratory behavior, whereas in smaller rooms, less variation in position or viewing angle is observed.

5 CONCLUSION

In this paper, we analysed the user behaviour during a photorealistic telepresence experiment developed on a volumetric social VR system. We mainly investigated through objective metrics the influence of narrative elements of the story, such as dialogues or interactive task, on participants' movements. Our results show indeed that the motion during the VR experience was affected by the storytelling. More static and focused behaviour happened when a task (either to switch on the light or press a button in a controller) was requested to be done by a specific participant. On the contrary, exploration movements were more frequent when virtual characters were talking in the scene. These observations are key factors to be further investigated, in order to develop immersive and interactive applications. Further work is needed to understand the effect of storytelling elements and cues on user behaviour also from a subjective point of view, in order to design social VR experiences that can effectively be optimized around the users.

REFERENCES

[1] E. Alexiou, N. Yang, and T. Ebrahimi. 2020. PointXR: A toolbox for visualization and subjective evaluation of point clouds in virtual reality. (2020).

[2] F. Buttussi and L. Chittaro. 2017. Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE transactions on visualization and computer graphics* (2017).

[3] C. Christou, A. Tzanavari, K. Herakleous, and C. Poullis. 2016. Navigation in virtual reality: Comparison of gaze-directed and pointing motion control. In *18th mediterranean electrotechnical conference*.

[4] X. Corbillon, F. De Simone, and G. Simon. 2017. 360-degree video head movement dataset. (2017).

[5] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (May 2007), 175–191.

[6] Ronald Aylmer Fisher et al. 1934. Statistical methods for research workers. *Statistical methods for research workers.* 5th Ed (1934).

[7] G. Freeman and D. Maloney. 2021. Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality. (2021).

[8] G. Gamelin, A. Chellali, S. Cheikh, A. Ricca, C. Dumas, and S. Otmane. 2020. Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. *Personal and Ubiquitous Computing* (2020).

[9] J. Jansen, S. Subramanyam, R. Bouqueau, G. Cernigliaro, M. M. Cabré, F. Pérez, and P. Cesar. 2020. A Pipeline for Multiparty Volumetric Video Conferencing: Transmission of Point Clouds over Low Latency DASH. In *Proceedings of the 11th ACM Multimedia Systems Conference*.

[10] James T Kost and Michael P McDermott. 2002. Combining dependent P-values. *Statistics & Probability Letters* 60, 2 (2002), 183–190.

[11] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. 2017. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology.*

[12] D. A. Le, B. MacIntyre, and J. Outlaw. 2020. Enhancing the Experience of Virtual Conferences in Social Virtual Environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW).* IEEE, 485–494.

[13] J. Li, V. Vinayagamoorthy, R. Schwartz, W. IJsselsteijn, D. A. Shamma, and P. Cesar. 2020. Social VR: A New Medium for Remote Communication and Collaboration. In *Extended Abstracts of the 2020 ACM CHI Conference on Human Factors in Computing Systems.*

[14] J. McVeigh-Schultz, A. Kolesnichenko, and K. Isbister. 2019. Shaping Pro-Social Interaction in VR: An Emerging Design Framework. In *Proceedings of the 2019 ACM CHI Conference on Human Factors in Computing Systems.*

[15] F. Moustafa and A. Steed. 2018. A longitudinal study of small group interaction in social virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology.*

[16] A. T. Nasrabadi, A. Samiei, A. Mahzari, R. P. McMahan, R. Prakash, M. C. Q. Farias, and M. M. Carvalho. 2019. A Taxonomy and Dataset for 360° Videos. Association for Computing Machinery.

[17] A. Revilla, S. Zamarvide, I. Lacosta, F. Perez, J. Lajara, B. Kevelham, V. Juillard, B. Rochat, M. Drocco, N. Devaud, O. Barbeau, C Charbonnier, P. de Lange, J. Li, Y. Mei, K. Ławicka, J. Jansen, I. Reimat, S Subramanyam, and P. Cesar. 2021. A Collaborative VR Murder Mystery using Photorealistic User Representation. In *Proceedings of the IEEE Virtual Reality Conference.*

[18] S. Rossi, F. De Simone, P. Frossard, and L. Toni. 2019. Spherical clustering of users navigating 360 content. In *IEEE International Conference on Acoustics, Speech and Signal Processing.*

[19] S. Rossi, C. Ozcinar, A. Smolic, and L. Toni. 2020. Do Users Behave Similarly in VR? Investigation of the User Influence on the System Design. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* (2020).

[20] S. Rossi and L. Toni. 2020. Understanding user navigation in immersive experience: an information-theoretic analysis. In *Proceedings of the 12th ACM International Workshop on Immersive Mixed and Virtual Environment Systems.*

[21] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li, et al. 2018. Emerging MPEG standards for point cloud compression. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* (2018).

[22] V. Sitzmann, A. Serrano, A. Pavel, M. Agrawala, D. Gutierrez, B. Masia, and G. Wetzstein. 2018. Saliency in VR: How Do People Explore Virtual Environments?. In *IEEE Transactions on Visualization and Computer Graphics.*

[23] S. Subramanyam, J. Li, I. Viola, and P. Cesar. 2020. Comparing the Quality of Highly Realistic Digital Humans in 3DoF and 6DoF: A Volumetric Video Case Study. In *IEEE Conference on Virtual Reality and 3D User Interfaces.*

[24] S. Subramanyam, I. Viola, A. Hanjalic, and P. Cesar. 2020. User Centered Adaptive Streaming of Dynamic Point Clouds with Low Complexity Tiling. In *Proceedings of the 28th ACM International Conference on Multimedia.*

[25] C. Swindells, B. A Po, I. Hajshirmohammadi, B. Corrie, J. Dill, B. Fisher, and K. Booth. 2004. Comparing CAVE, wall, and desktop displays for navigation and wayfinding in complex 3D models. In *IEEE Proceedings Computer Graphics International.*

[26] J. Van der Hooft, T. Wauters, F. De Turck, C. Timmerer, and H. Hellwagner. 2019. Towards 6-DoF HTTP adaptive streaming through point cloud compression. In *Proceedings of the 27th ACM International Conference on Multimedia.*

[27] J. Williamson, J. Li, V. Vinayagamoorthy, D. A Shamma, and P. Cesar. 2021. Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop. *arXiv preprint arXiv:2101.05300* (2021).