Eurographics Workshop on Intelligent Cinematography and Editing (2022)
R. Ronfard and H.-Y. Wu (Editors)

(Re-)Framing Virtual Reality

Rémi Sagot-Duvauroux[1] François Garnier[1] and Rémi Ronfard[1,2]

[1]ENSAD, PSL University, EnsadLab, Spatial Media, Paris, France
[2]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France

Abstract

We address the problem of translating the rich vocabulary of cinematographic shots elaborated in classic films for use in virtual reality. Using a classic scene from Alfred Hitchcock's "North by Northwest", we describe a series of artistic experiments attempting to enter "inside the movie" in various conditions and report on the challenges facing the film director in this task. For the case of room-scale VR, we suggest that the absence of the visual frame of the screen can be usefully replaced by the spatial frame of the physical room where the experience takes place. This "re-framing" opens new directions for creative film directing in virtual reality.

CCS Concepts

• ***Applied computing*** → ***Media arts;***

1. Introduction

In their survey of cinematic virtual reality movie-making or "cine VR", Williams et al. [WLL21], exclude the case where the audience is free to move in the virtual world. In contrast, authors such as Nonny de la Peña believe that the future of cinema is in "fully embodied, walk around, room-scale, volumetric experiences" [Buc19], in short "room-scale" VR. This work is part of ongoing art & science, practice-based research project investigating montage as a narrative vector for room-scale VR [RSDR22].

Room-scale VR presents special challenges for intelligent cinematography and editing. First of all, cinematography is almost entirely under the control of the player. The film director only gains control of the camera position at each shot transition, where he can decide where the immersant is positioned relative to the virtual world. Immediately after that, the immersant regains the full control of its movement in the room, which results in unpredictable "camera movements" and "framings".

In this communication, we focus on one important aspect of room-scale VR cinematography and film editing, the absence of the frame. The frame, according to Sergueï Eisenstein, is a "montage cell" [Eis76]. What would then be a "montage cell" in VR if we don't have a frame? One of the most important distinctions in the traditional grammar of cinema is the division of shot sizes, which refer to the screen size of the subjects, relative to the frame size. In the absence of a screen and a frame, what happens to long shots, medium shots and close shots? Does the classification become obsolete? Should it by replaced?

In room-scale VR, the perceptual boundaries of a continuous virtual reality experience are based on the subject's freedom of movement in the real environment. First of all, the context of the experience (museum, arcade room, dedicated place, festival, home...) implies a specific reception state leading to a first "frame" on the immersant's movements. Then, in a more concrete way, the technical equipment and the perimeter of the physical space tracked by the device therefore seems to correspond to a frame. In the virtual environment, it defines a scaled, spatialized and oriented zone in which the subject acquires a limited walking and acting area. We therefore decide to take this zone of potential action as our editing cell. The frame in virtual reality no longer delimits a portion of the image as in cinema, but a zone of possible actions in a space.

In Section 2, we review previous work on this topic in both cinematic VR and room-scale VR. Then in Section 3, we outline our contributions. We present our experimental set-up for experiencing a movie "from the inside" in VR in Section 4. We built a synthetic animated 3D scene reproducing the well known crop-duster sequence in "North by Northwest" by Alfred Hitchcock. This sequence is nine minute long, mostly silent, and contains 133 shots showing a variety of shot sizes and compositions. Our goal is to reproduce the decoupage of the sequence and to experience it "inside the movie".

In Section 5, we present three steps of our research-creation process towards this goal. In the first step, we watched the movie in a VR headset at a fixed position and orientation, as if watching the movie on a very large screen. In the second step, we allowed ourselves to orientate our head in all directions. In the third step, we could move around within the limits of a small room. Our main finding is that the camera positions used by Hitchcock in the film version need to be redesigned completely in all three versions. We also find that the best shot choices in cinematic and room-scale VR

DOI: 10.2312/wiced.20221051

Figure 1: *Equirectangular 360° view of our version of Alfred Hitchcock's "North by Northwest" in VR. In the absence of a visual frame, we can impose a spatial frame around the immersant, with its own affordances for action, which gives rise to a new vocabulary of shot types.*

are quite different. In the case of room-scale VR, we find that the missing "visual frame" in the field of view is replaced by the "spatial frame" of the room where the experience takes place.

Finally in Section 6, we discuss how this new taxonomy of VR "shots" makes it possible to "frame" the experience of the immersant, by choosing to give him more or less space, and we conjecture that this new method of framing can be a powerful tool for VR film makers and opens new directions for VR cinematography and film editing.

2. Related work

Framing in cinematic VR was investigated in several experimental "studies" by Naimark [Nai16], where he observed a limited range of "useful" distance to subjects around a panoramic camera between one meter for close-ups and five meters for long shots. Naimark observed that actors further than five meters from the camera were hardly recognizable. This hints at a very limited range of "useful" shots sizes, compared to traditional cinematography.

On a more conceptual level, Mateer examines how the traditional craft of film directing can be adapted to the special case of cinematic VR [Mat17] and argues that the classical continuity style of film directing is applicable to cinematic VR, and can be used to predict and control the user's viewpoint within the virtual scene, based on a careful interpretation of the story elements present in the screenplay. Mateer grounds his analysis in transportation theory [GBK04], where the goal of the narrator is to transport the reader into the story world.

Gödde et al. review the state of the art in cinematic narration in VR and propose novel direction for research, taking into account the spatial and temporal story density in a VR film [GGSB18]. They propose a re-interpretation of framing and editing for the case of cinematic VR. Tricart provides a comprehensive overview of virtual reality film-making, including game engine-based VR which "opens up a lot more possibilities ranging from branching storytelling to room-scale games" [Tri17]. In this paper, we focus on directing techniques suitable for building "room-scale games" and more specifically "room-scale movies". Garnier [Gar21] provides a detailed accounts of the various geometries involved in viewing film, either on a traditional screen or in a virtual reality headset, and emphasizes the importance of a proxemic interpretation of shot sizes in film. From another perspective, Pope also recommends the use of proxemics for staging in VR following some well established theatre techniques [PDSS17].

Closer to our concern, Rothe et al. propose an analysis of camera control in cinematic virtual reality [RSH19] and relate shot sizes and camera distances in cinematic VR under the framework of proxemics. Rothe also examines the effect of camera height in distance and size perception in cinematic virtual reality [RKAH18]. In our work, we extend those findings for the case of room-scale VR, where the immersant is also free to move in the virtual world. In this case, distance is not measured only in terms of apparent visual size, but also in motion parallax and capacity of action.

3. Contributions

Most of the previous work in virtual reality cinematography has focused on cinematic VR, rather than room-scale VR. We believe a possible reason for this is the lack of suitable data sets that can be used to perform experiments and conduct studies in room-scale VR. In the case of cinematic VR, a large corpus of 360 degree videos can be used. Building a data set for room-scale VR is comparatively more challenging because a room-scale "movie" is in fact a

real-time 3D game. In order to build experiments, it is therefore necessary to create all the necessary 3D assets and animations, and organize them into a real-time game.

The first contribution of our paper is a new data set reproducing the "crop duster" scene in Alfred Hitchcock's movie, "North by Northwest". The data set contains all necessary assets, animations and cameras to approximately reproduce the movie scene in rough layout animation. The data set is suitable for performing experiments in room-scale VR film making, e.g. by changing the camera angles and the ordering and duration of shots.

The second contribution of our paper is a principled investigation of what it means to experience a classic movie in VR, and how the camera angles and the ordering of shots need to be modified from the original version. Based on our ongoing art-based research, we use those experiment to approach the possibility of montage in room-scale VR. We make the strong hypothesis that room-scale VR requires a different "shot grammar" from both traditional film and cinematic VR. We are guided by the observation that the sense of movement [Ber02] guides our perception of space at least as much as the sense of vision when we dynamically explore a virtual world around us. Motion parallax also plays a major role. As a result, the scale of a shot cannot simply be measured by the distance between the immersant and the subject, or the apparent size of the subject in the immersant's field of view.

To investigate those issues, we attempted to re-create the experience of a movie with many different shot sizes in three different conditions (Fig. 2). In the first condition, we were standing, but we had to keep our heads fixed while looking straight ahead, as the camera of Hitchcock. This setup was close to the reception state of a stereoscopic film with a wide field of view. In the second condition, we allowed ourselves to turn our heads through 360 degrees establishing the reception state of a stereoscopic cinematic VR experience with three degrees of freedom (rotation of the head in the x, y and z axis). In the third condition, we could move freely through the physical room tracked by the system with six degrees of freedom (rotation and translation of the head in the x, y and z axis).

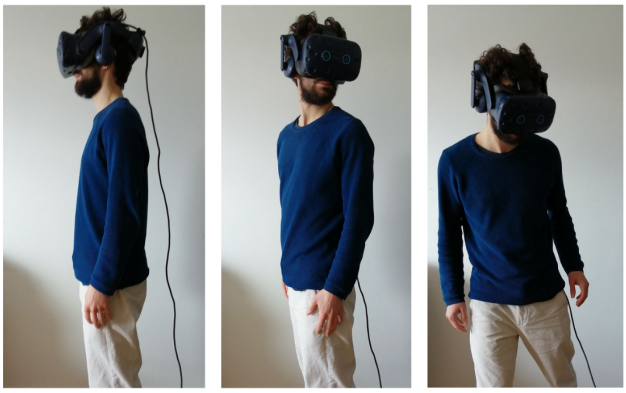

Figure 2: *Immersant's reception states with the three experimental conditions. Left: fixed head and body. Center: free head and fixed body. Right: free head and body.*

Previous work has emphasized the absence of a frame in cinematic VR, and proposed solutions for directing movies without a frame. The third contribution of our work is the finding that there *is* a frame in room-scale VR, but of a very different kind - it is the physical volume of the room where the experience is taking place. This volume constrains the movements that the immersant can perform. By changing the relative sizes of the virtual and the physical worlds, we can define "wide shots" with miniature non player characters where the room appears to be 100 meters wide; "long shots" with dolly-size characters, where the room appears to be 10 meters wide; "medium shots" with three-quarter-size virtual characters; "point of view shots" with real-size virtual characters; and even "close shots" with larger-then-life virtual characters where movements are limited to a meter or less.

4. Movie scene reconstruction

Previous work in intelligent cinematography and editing has used 3D reconstruction of movie scenes for purposes of experimentation and validation. A detailed reconstruction of the cafe scene in "Back to the future" was used by Galvane et al. [GRLC15] for demonstrating film editing, and was also used to build previz in VR. For our purpose, the cafe scene is not entirely adequate because it consists mostly of medium shots.

We instead decided to reconstruct another famous movie scene, showing a much larger variety of shot sizes and shot compositions. The crop-duster sequence in "North by Northwest" is one of the most widely studied cinematic sequences in the history of cinema. Raymond Bellour included a comprehensive shot by shot analysis in his classic book on "The analysis of film" [Bel00]. He noted that this sequence "is constructed in its entirety according to the alternation seeing/seen between the subject and the object of his vision; this alternation is regularly interrupted by a coming together in a single field of the subject and the object, according to a very complex hierarchization of repetitions and differences, all of this being continued for 133 shots" [Ber79]. Whether such a pattern can be reproduced in VR is an interesting question in its own right. The sequence contains a variety of shot sizes, ranging from the opening wide shot, to many alternating long shots and medium shots, and a smaller number of medium close-ups and one close up of the main character, played by Cary Grant. There is very little camera motion in those shots, which makes the sequence appropriate for viewing in a virtual reality headset. Interestingly, the rhythm of the sequence is largely imposed by the editing, which is moderately fast, with an average shot length of 4 seconds.

We manually reconstructed the entire set and all the props, characters and vehicles present in the scene using representative keyframes from the movie. We then created a rough "layout" animation of all the character and vehicle movements, synchronized frame by frame with the original movie. We used a floor plan view diagram from Bellour's book as a reference for approximately reconstructing the trajectories of all characters, vehicles and the plane, as shown in Fig. 3, then we fine-tuned them to match the existing views in the movies. Some movements could not be seen in the original shots and had to be crafted manually to plausibly match the visible parts.

To compute the camera angles, we used a diagram attributed to

cinematographer Robert Burks and reproduced in [SG07]. The diagram contains 61 numbered cameras, which can be traced back to the 133 actual shots in the sequence (more than one shot can use the same camera). Using this diagram as a reference, we manually placed and oriented the cameras in the 3D scene to reproduce the camera angles as closely as possible. To control the shot compositions, we used the "prose storyboard" annotation of the scene by [RGBM22], which contains a complete breakdown of all 209 compositions in the sequence, taking into account camera and actor movements.

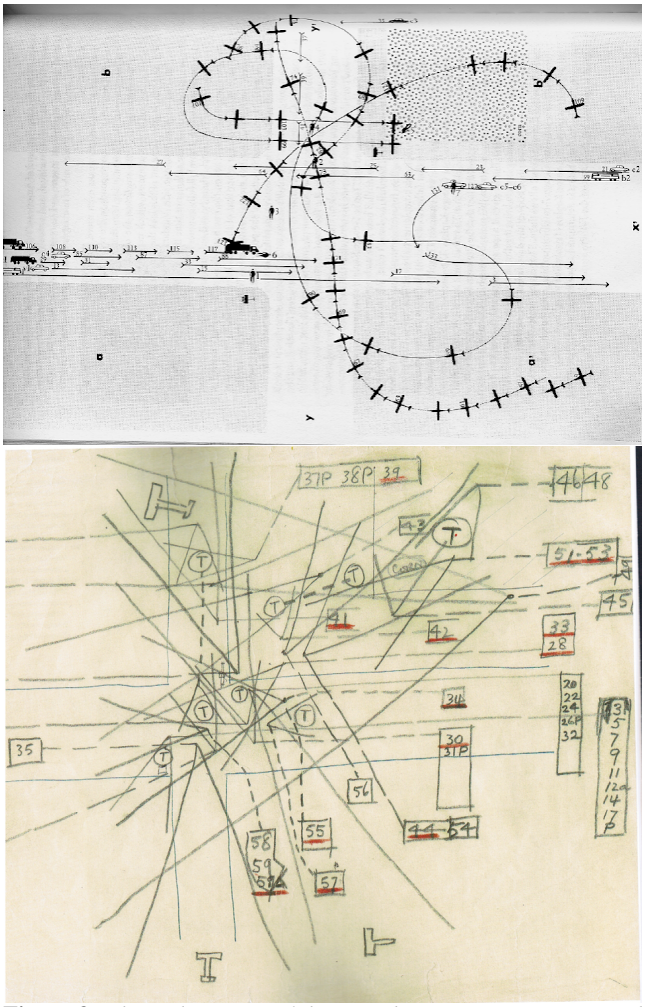

Figure 3: *Floor plan views of the crop duster scene reconstructed from the movie. Top: Trajectories of all characters, vehicles and plane (reproduced from [Bel00]). Bottom: All 61 camera angles in the scene, as notated by cinematographer Robert Burks (reproduced from [Hig94]).*

5. Artistic experiments

We now describe three experiments in re-framing and re-editing the crop-duster sequence. They are not scientific experiments per se, but rather "artistic experiments" guided by the practice of the first author as a professional film editor. We used a research-creation methodology based on the theorising by an artist of his own ongoing creation process. We investigated three reception conditions (Fig. 2) specific to immersion with a virtual reality headset: cinematic VR with fixed head and body, cinematic VR with free head and fixed body, and room-scale VR with free head and body. We provide an explanation of the strategy we used in each case for guiding the immersant from "shot" to "shot". We also describe our findings, which are mostly based on the subjective evaluations of the three authors. We recorded one session per experiment, and they are included as supplementary material in the accompanying video, which can be found at `https://vimeo.com/681905742` for reference and comparison.

5.1. Cinematic VR with fixed head and body

In our first experiment, we edited the immersant's points of view according to the 133 camera positions computed from the movie scene. Reproducing the exact timeline of the movie, we asked ourselves not to move watching straight like Hitchcock's camera. Several problems emerged.

Firstly, we observed that when the camera is tilted and the ground plane is not horizontal, we tend to feel a vertigo that can lead to motion sickness. We also experienced motion sickness in the few instances where the camera moves along a dolly track, especially in cases where the camera is tilted down. Otherwise, we noticed that backward camera movements ("dolly out") were subjectively preferable and easier to handle than forward camera movements ("dolly in").

Then, we observed that the rhythm of the montage seemed faster than when watching the original movie. This can be explained by the longer time needed for processing stereoscopic images with a large field of vision. Moreover, the depth of field is determined by the eyes of the immersant who chooses his/her point of focus.

We also subjectively observed that the aesthetic impact of Hitchcock's framing choices where lost in a majority of cases. The image compositions didn't provoke the same artistic impressions when watching them in the virtual environment. The large field of view and the absence of imposed depth of field encourage our eyes to wander across the perceived space instead of focusing on the characters. An eye tracking analysis would be useful for determining how the immersant's focus evolve. Furthermore, we felt that the distance between the camera and its filmed subject chosen by Hitchcock did not work well in this experiment. Most of the time, we felt too close or too far from the subject.

Finally, we felt that the cuts looked unmotivated, arbitrary and meaningless. Whereas in Hitchcock's film, the montage helps to build suspense by loading each shot with dramatic tension, our experiment completely loses this strength, leaving only a succession of disconnected viewpoints.

5.2. Cinematic VR with free head and fixed body

In our second experiment, we allowed ourselves to turn the head at 360° to be in the same reception state of a cinematic virtual reality experience. Camera positions are no longer seen as fixed vectors

but as zones proposing a multitude of points of view from a fixed point.

In order to reduce the motion sickness issues of the first experiment, we imposed that the zones had to always be parallel to the ground. In other words, they could not be tilted in their x or z axis. As we were able to look around, our reception contract changed with the narrative. We could now choose where to look and where not to look. With this new freedom, the cuts became even more problematic. Firstly, because they broke our engagement in the reading of the scene around us, they became frustrating. Secondly, we observed that cuts often led to disorientation bringing the immersant back to his/her own orientation in space rather than to the narrative. The cuts become teleportations and broke the continuity of the narrative.

As expected, the rhythm of the montage in this case felt even faster and meaningless. For example, at the beginning of the sequence, Hitchcock uses shots of about three seconds alternating between a third person view of the character looking around and a first-person view of the landscape around him. Reproducing this editing sequence in a set up where we can choose where to look does not work well. Instead, the immersant's rotating gaze should be taken into account to produce the same effect.

5.3. Room-scale VR with free head and body

In our third experiment, we allowed ourselves to move through the virtual space relatively to the physical space available. Our editing cells evolved from fixed 360° camera positions to oriented areas to walk in. In this case, we also adapted the montage of the sequence regarding the issues we observed in the two previous experiments.

We first experimented with oriented areas ("spatial frames") of the same "size" in the virtual world. In other words, we kept the virtual world at the same scale for the entire sequence. With this constraint, we discovered that there were not a significant difference with the 360° video reception regarding the aesthetic impact of the montage. As we always kept the same size in the environment, we felt to teleport in different points of view, but without changing the relationship between the represented scene and our narrative distance to it.

We then decided to explore how to adapt Hitchcock's shot scales into different immersant's scale in relation to the environment. According to the scale of our virtual body, the representation space is not perceived with the same density. We have the intuition that it is the variations of this density which will reproduce the distance effects of cinematographic shot scale changes.

To translate the effect of a wide shot, we gave the immersant the size of a giant watching the scene in an allocentric reference frame. As seen in Fig. 4, the giant point of view not only enables us to have a high point of view without having the effect of vertigo, but also densifies the distance between us and the elements of the scene. We do not need to move a lot to get close to an element or another. Thus, this scale approaches the impression of a cinematographic wide shot allowing us to embrace globally a broad space to locate a scene.

The second shot of the sequence is a long shot that shows the

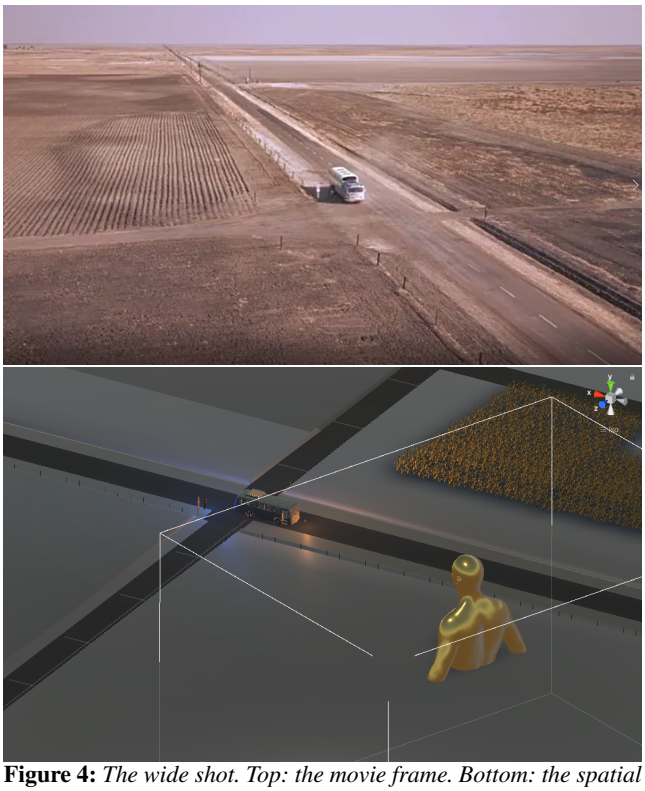

Figure 4: *The wide shot. Top: the movie frame. Bottom: the spatial frame.*

character more distinctly within his environment. We therefore had to downscale the immersant to reduce the gap between his/her size and the one of the character (as seen in Fig. 5). Thus, we could start to see him as a little toy within which we can start to project in a third person point of view.

For the full shot, we had to reduce the distance further in order to feel closer to the character's density of action. As shown in Fig. 6, this causes the character to grow from the size of a small toy to that of a large doll. In this way, we obtain the effect of perceiving the character from head to toe, being able to focus more significantly on his actions, while at the same time having a significant scale difference to translate the distance effect of Hitchcock's full shot.

The medium shot must create the effect of a certain intimacy with the character, making him more accessible, and letting his vulnerability show through. To translate this effect into a scale relationship between the immersant and the character, we first tried to give the immersant the same scale as the character, putting for the first time his/her feet on the ground. At this scale, we realized that suddenly, we inhabited the space of representation. Having the same size of the character made us part of the environment and broke the distance effect between the representation and us. To avoid this effect we upscaled and lowered the immersant in order to align his/her head to the one of the character (as seen in Fig. 7). In this way, we could achieve the distance effect of the medium shot without feeling to be part of the diegesis.

In Hitchcock's sequence, the montage alternates frequently between the main character and his point of view. We said before that

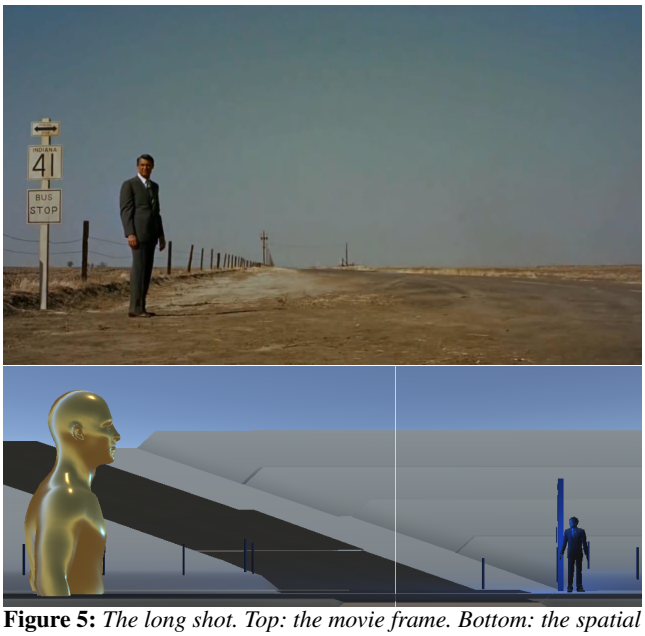

Figure 5: *The long shot. Top: the movie frame. Bottom: the spatial frame.*

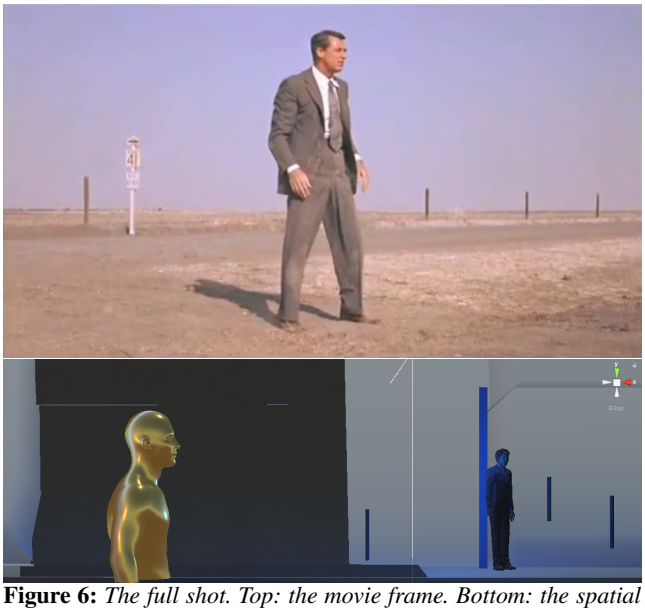

Figure 6: *The full shot. Top: the movie frame. Bottom: the spatial frame.*

this switching could be made by the rotation of the immersant's body and gaze. But do these rotations reproduce the narrative effects of the montage, which is to change our reception between empathy (putting ourselves in the other person's shoes) and sympathy (understanding the other person from our own perspective)? In order to reproduce the reception state of the character's point of view, we decided to give the immersant the same size as the character in this case, having the ground under his feet and perceiving the world at the same scale (as seen in Fig. 8).

Finally, came the question of the close-up. How can we reproduce in VR the complex and strange effect of a cinematic close-up that brings the camera point of view closer to a subject than we naturally would in real life? To reach this effect, we decided to downscale the immersant and make him/her much smaller than the character (as seen in Fig. 9). One significant issue of this small size within the environment, is that translations of the body have much less amplitude to explore the scene. On the other hand, as the whole world is now bigger, our head rotations have much more impact in the changing of what we see in our field of vision. In other words, the more we reduce our virtual body size relatively to the environment, the more we are confined in a located area but the more we have to rotate our head in order to follow salient elements of the narrative.

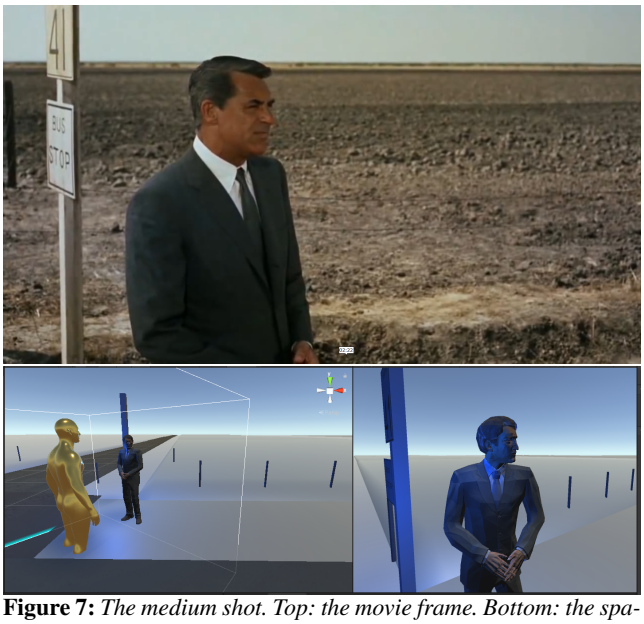

Figure 7: *The medium shot. Top: the movie frame. Bottom: the spatial frame.*

Our new definition of shot sizes is an aesthetic proposal which makes it possible to also redefine the montage of the sequence, in a way that is compatible with the movements of the immersant's head and body. As we noticed earlier, the fast pace of editing in Hitchcock's version creates an unmotivated and ineffective montage when reproduced exactly shot by shot. Our new definition of shot sizes allows to respect the alternation of different shot sizes (in spatial terms) in Hitchcock's movie, while at the same time giving more time to the immersant to explore the spatial frame in each case. As can be seen in the accompanying video, this leads us to change the shot durations and propose a slower montage that retains the rhythm, the tension, the character's emotion and the right distance with the narrative of Hitchcock's sequence.

Our different scales involved an adaptation of the rhythm of the montage regarding their aim in relation to the ongoing actions of the sequence. According to the richness of information present in the action area, we had to take into account the longer processing time of the spatial perception. We saw that each scale implies a different engagement of the immersant regarding its body movements and rotations. This bodily engagement induces a higher agency in

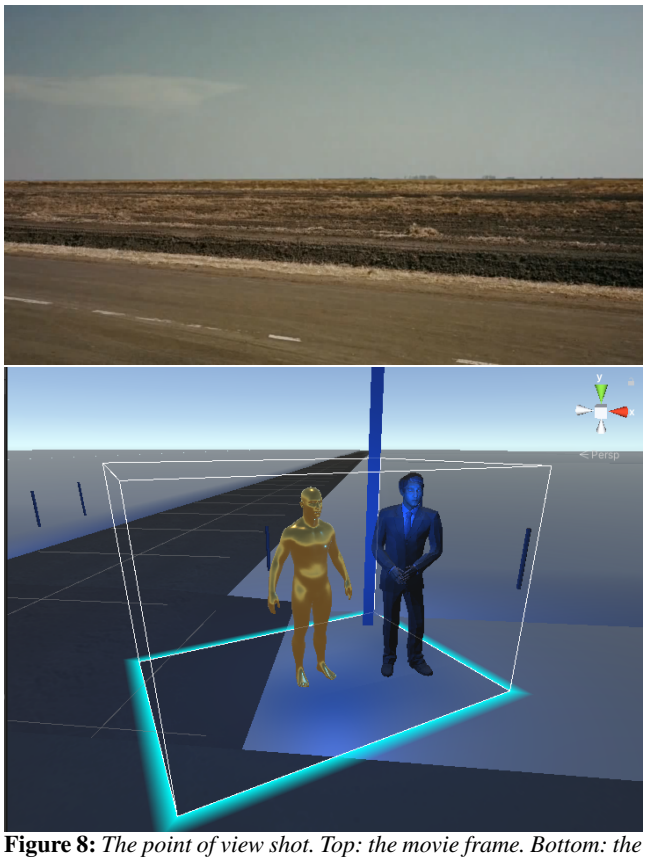

Figure 8: *The point of view shot. Top: the movie frame. Bottom: the spatial frame.*

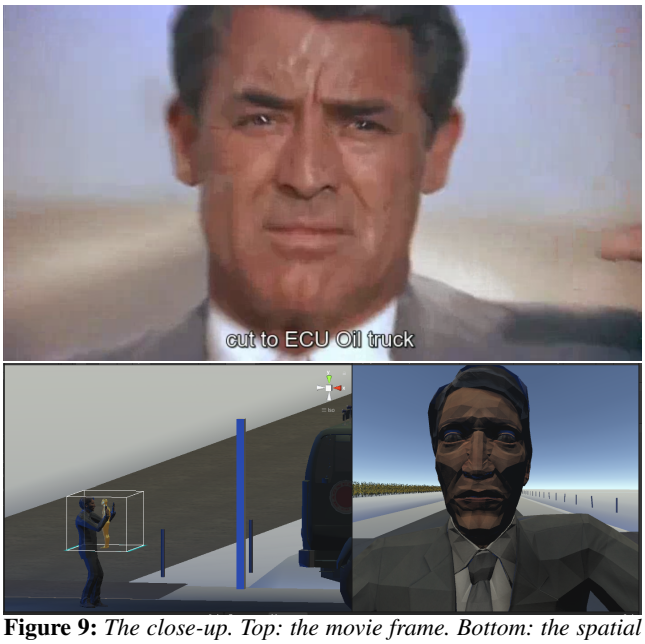

Figure 9: *The close-up. Top: the movie frame. Bottom: the spatial frame.*

the reception and requires a rethinking of the duration of the editing cells according to their new potentialities of action. The longer the cell, the more control the immersant has over his/her perception of the scene and the more frustrating it can be to be pulled back to the point of view imposed by the editing.

Writing about their Kino-eye in the 1920's, Dziga Vertov and the Kinoks stated that it was the intervals (between two movements) that constitute the material that "drives the action towards the kinetic denouement" [Ver18]. If we follow this theory, the quality of the editing in virtual reality should be highly impacted by the quality of the movement of the person experiencing it. The kinetic power of the montage would then be proportional to the energy deployed by the viewer's movements. In our frame taxonomy, these potential movements depend of the immersant's scale that brings different densities of action. It therefore seems interesting to take into account these variations of deployed energy in our artistic practice of editing the scale discontinuities in time and space.

6. Discussion and future work

The new shot scale that we propose in this paper is part of an investigation of the narrative possibilities of room-scale VR. Until recently, a majority of movies made for cinematic and room-scale VR has focused on real-time, immersive experiences where the time of the story is more or less equal to the actual time of the experience and the space of the story world is more or less equal to the physical space of the room. This appears to limit the use of VR to "Aristotelian" drama that obey the "classical unities" of time, space and action.

Previous work has investigated the possibility of montage in virtual reality [SSRB*17, COFS18, KMR18, Hod19, PV19, MGS20]. In "North by Northwest", the space of the story world is 100 times larger than the space of the room, and the time of the story is significantly longer than the nine minutes of the experience. Changing the shot sizes appears to be a key element in maintaining the high pace of editing necessary to convey such a complex story. Together with the above previous work, this seems to indicate that varying the sizes of the "spatial frames" in a virtual story world can be a powerful storytelling device for virtual reality film makers. As a limitation, the artistic experiments reported in this work cannot claim to be scientific experiments, and our preliminary findings will need to be confirmed in future work with a proper user study on a larger amount of subjects, using appropriate qualitative evaluation protocols.

Based on the preliminary findings in this paper, we are building an experimental film editing workbench, where we use the wide shot to control the staging and framing of a room-scale VR movie by placing room-scale "frames" of different sizes in the story world, and defining "cuts" between them. Cuts can further be triggered by "cues" such as predefined movements of the immersant or non player characters, dialogues spoken by the non player characters, or elapsed time in any given shot. With this new vocabulary of spatial "shots" and "cuts", we hope to construct more interesting and immersive room-scale VR movies that develop complex stories both in time and in space.

7. Conclusion

In this paper, we hypothesized that in order to translate the effects of cinematic shot scale changes in the case of room-scale virtual reality experiences, we needed to rethink the way of framing the actions. Thanks to a 3D animation reproduction of a sequence from Alfred Hitchcock's "North by Northwest", we were able to confront the original editing intended for reception on a 2D screen with the phenomenological particularities of three reception conditions specific to immersion with a virtual reality headset (fixed head and body; free head and fixed body; free head and body).

The two first creative experiments led us to different conclusions regarding the adaptability and the way to rethink the aesthetic effects of shot scales within a room-scale virtual reality reception. We first saw that positioning the viewer's gaze at the same position as Hitchcock's camera did not reproduce the same effects regarding the focus and distance toward the character. We then observed that by releasing the immersant's head rotation, the original montage became even less meaningful, and too fast. We also felt that cuts were more felt as teleportations than a meaningful scale change.

Our third experiment enabled us to put to test our hypothesis of a "spatial frame" around the immersant, leading to a redefinition of shot sizes in room-scale virtual reality. By testing a variety of scales of the immersant and his/her physical tracking space relative to the scale of the virtual environment, we experimentally found different size ratios that seem to reproduce the distance effects of the cinematic frame. Moreover, the closer we get to the character's scale, the more we come to inhabit his space and the more we feel we belong to the diegesis. Finally, we analyzed that each scale brings a difference of amplitude of our bodily action in the virtual environment (a more or less large space is reachable) implying different agencies in the reception and therefore a new rhythm to find for the montage of these scale variations in space and time.

References

[Bel00] BELLOUR R.: *The Analysis of Film*. Indiana University Press, 2000. 3, 4

[Ber79] BERGSTROM J.: Alternation, segmentation, hypnosis: Interview with raymond bellour. *Camera Obscura* (1979), 70–103. 3

[Ber02] BERTHOZ A.: *The Brain's Sense of Movement*. Harvard University Press, 2002. 3

[Buc19] BUCKMASTER L.: What films will be like in 20 years. https://www.bbc.com/culture/article/20190508-what-will-films-be-like-in-20-years, 2019. 1

[COFS18] C. O. FEARGHAIL C. OZCINAR S. K., SMOLIC A.: Director's cut - analysis of aspects of interactive storytelling for vr films. In *International Conference for Interactive Digital Storytelling (ICIDS)* (2018). 7

[Eis76] EISENSTEIN S.: *Le film: sa forme, son sens*. C. Bourgois, 1976. 1

[Gar21] GARNIER F.: Artistic practices in digital space: An art of the geometries of movement? In *Space-Time Geometries for Motion and Perception in the Brain and the Arts*. Springer, 2021. 2

[GBK04] GREEN M., BROCK T., KAUFMAN G.: Understanding media enjoyment: the role of transportation into narrative worlds. *Communication Theory 14*, 4 (2004), 311–327. 2

[GGSB18] GÖDDE M., GABLER F., SIEGMUND D., BRAUN A.: Cinematic narration in vr - rethinking film conventions for 360 degrees. In *HCI* (2018). 2

[GRLC15] GALVANE Q., RONFARD R., LINO C., CHRISTIE M.: Continuity editing for 3d animation. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence* (2015), AAAI'15, AAAI Press, p. 753–761. 3

[Hig94] HIGGINS S. C.: *The moviemaker's workspace: towards a 3D environment for pre-vizualization*. Master's thesis, Massachusetts Institute of Technology, 1994. 4

[Hod19] HODGKINSON G.: Cut, don't cut - moving the viewer through a story in animated vr. In *Proceedings of 7th International Conference on Illustration and Animation (CONFIA 2019)* (2019), p. 50-57. 7

[KMR18] K. MOGHADAM C. B., RAGAN E.: Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness. *IEEE Transactions on Visualization and Computer Graphics* (2018). 7

[Mat17] MATEER J.: Directing for cinematic virtual reality : how traditional film director's craft applies to immersive environments and notions of presence. *Journal of Media Practice 18* (05 2017). 2

[MGS20] MARAÑES C., GUTIERREZ D., SERRANO A.: Exploring the impact of 360° movie cuts in users' attention. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2020), pp. 73–82. 7

[Nai16] NAIMARK M.: VR cinematography studies for google. http://michaelnaimark.medium.com/vr-cinematography-studies-for-google-8a2681317b3, 2016. 2

[PDSS17] POPE V. C., DAWES R., SCHWEIGER F., SHEIKH A.: The geometry of storytelling: Theatrical use of space for 360-degree videos and virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017), CHI '17, p. 4468–4478. 2

[PV19] PILLAI J. S., VERMA M.: Grammar of vr storytelling: Narrative immersion and experiential fidelity in vr cinema. In *International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '19)* (2019). 7

[RGBM22] RONFARD R., GANDHI V., BOIRON L., MURUKUTLA A.: The prose storyboard language: A tool for annotating and directing movies (version 2.0, revised and illustrated edition). In *Eurographics Workshop on Intelligent Cinematography and Editing* (2022). 4

[RKAH18] ROTHE S., KEGELES B., ALLARY M., HUSSMANN H.: The impact of camera height in cinematic virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (2018), VRST '18. 2

[RSDR22] RÉMI SAGOT-DUVAUROUX RÉMI RÉMI RONFARD F. G.: Montage as a narrative vector for virtual reality experiences. In *ConVRgence (VRIC) Virtual Reality International Conference Proceedings* (2022). 1

[RSH19] ROTHE S., SARAKIOTIS V., HUSSMANN H.: Where to place the camera. In *25th ACM Symposium on Virtual Reality Software and Technology* (2019), VRST '19. 2

[SG07] SCHMENNER W., GRANOF C.: *Casting a Shadow: Creating the Alfred Hitchcock Film*. Northwestern University Press, 2007. 4

[SSRB*17] SERRANO A., SITZMANN V., RUIZ-BORAU J., WETZSTEIN G., GUTIERREZ D., MASIA B.: Movie editing and cognitive event segmentation in virtual reality video. *ACM Trans. Graph. 36*, 4 (2017). 7

[Tri17] TRICART C.: *Virtual Reality Filmmaking: Techniques & Best Practices for VR Filmmakers*. Routledge, 2017. 2

[Ver18] VERTOV D.: *Le ciné-oeil de la révolution, écrits sur le cinéma*. Les presses du réel, 2018. 7

[WLL21] WILLIAMS E. R., LOVE C., LOVE M.: *Virtual Reality Cinema: Narrative Tips and Techniques*. Routledge, 2021. 1