

Article

Viewer's Role and Viewer Interaction in Cinematic Virtual Reality

Lingwei Tong ^{1,*} , Robert W. Lindeman ¹  and Holger Regenbrecht ²¹ HIT Lab NZ, University of Canterbury, Christchurch 8041, New Zealand; gogo@hitlabnz.org² Department of Information Science, University of Otago, Dunedin 9016, New Zealand; holger.regenbrecht@otago.ac.nz

* Correspondence: kris.tong@pg.canterbury.ac.nz

Abstract: Cinematic Virtual Reality (CVR) is a form of immersive storytelling widely used to create engaging and enjoyable experiences. However, issues related to the Narrative Paradox and Fear of Missing Out (FOMO) can negatively affect the user experience. In this paper, we review the literature about designing CVR content with the consideration of the viewer's role in the story, the target scenario, and the level of viewer interaction, all aimed to resolve these issues. Based on our explorations, we propose a "Continuum of Interactivity" to explore appropriate spaces for creating CVR experiences to archive high levels of engagement and immersion. We also discuss two properties to consider when enabling interaction in CVR, the *depth of impact* and the *visibility*. We then propose the concept framework Adaptive Playback Control (APC), a machine-mediated narrative system with implicit user interaction and backstage authorial control. We focus on "swivel-chair" 360-degree video CVR with the aim of providing a framework of mediated CVR storytelling with interactivity. We target content creators who develop engaging CVR experiences for education, entertainment, and other applications without requiring professional knowledge in VR and immersive systems design.



Citation: Tong, L.; Lindeman, R.W.; Regenbrecht, H. Viewer's Role and Viewer Interaction in Cinematic Virtual Reality. *Computers* **2021**, *10*, 66. <https://doi.org/10.3390/computers10050066>

Academic Editors: Daniel Zielasko and Bernhard Riecke

Received: 30 March 2021
Accepted: 14 May 2021
Published: 18 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: cinematic VR; digital storytelling; interaction; 360-degree narrative

1. Introduction

The term Cinematic Virtual Reality (CVR) can be defined as a type of experience where the viewer watches omnidirectional movies using head-mounted displays (HMD) or other Virtual Reality (VR) devices [1]. Thus, the viewer can develop a feeling of being there within the scenes and can freely choose the viewing direction [2]. From a content point of view, we use the prefix "cinematic" or "narrative" to define those VR experiences that are narrative-based, instead of purely for novelty, entertainment, exploration, etc. A narrative virtual reality project can be a story-based drama, a documentary, or a hybrid production that features a beginning, middle, and an end. Its appearance may vary from simple 360-degree videos, where the only interaction for viewers is to choose where to look, to complex computer-generated experiences where the viewer can choose from multiple branches or even interact with objects and characters within the scene [3].

As CVR becomes more widely popular, content creators are trying to produce engaging narratives using this immersive medium. One of the major issues creators encounter is called the "Narrative Paradox", which is the tension between the user having freedom of choice and customization and the director controlling how the narrative plays out [4]. It poses a challenge for creators to balance engaging interactivity with dramatic progression. The second major issue linked to viewer control is that it can cause the viewer to miss important story elements, inducing a condition called "Fear of Missing Out" (FOMO) [5], and yielding weak narrative comprehension and low emotional engagement.

Many researchers have worked on solutions to these issues. We borrowed terms from a film director's skill set, namely the *Mise-en-scene* (a French term of where the things are in the scene), Cinematography, and Editing [6], to group them. We leave out the use of sound

as it is outside of the scope of this paper. Under *Mise-en-scene*, people have been trying to maintain narrative control and deliver dramatic experiences by changing the user's viewpoint in the scene (e.g., camera placement), the placement of action, and the story elements [7–9]. Other work comes from a cinematographic perspective and discusses how the spatial-temporal density of a story, the framing grammar, and editing techniques apply to CVR [3,10–12]. Finally, some researchers have explored methods to direct the viewer's attention within the immersive environment to important story elements [2,4,13,14]. These sit at the border between *Mise-en-scene* and Cinematography, as some use diegetic story elements in the scene, while others use visual cues and alternations.

We note, however, that these solutions have mainly focused on *the director's role*, as opposed to *viewer agency*. For example, CVR directors employ methods to direct the viewer's attention to important elements, compose a story script focused on balancing the spatial and temporal story density, and need to rethink framing and editing in both production and post-production stages. Little attention has been given to viewers, especially to the viewer's role and agency through interaction. We assume that this is because creators are most familiar with mature filmmaking techniques used in the industry, similar to how we have chosen our starting point stemmed from well-established terms. In traditional filmmaking, viewer agency is seldom brought into consideration, because in a cinema, the viewers are passively sitting and looking straight at the screen, with *zero interaction* with the content and no influence on the story. Thus, CVR creators take this as the given viewer scenario [6] and ignore viewer agency when migrating techniques to CVR. However, the viewing experience of swivel-chair CVR is quite different from traditional cinema. Firstly, although both are normally experienced while seated, a CVR viewer is able to turn her head or chair to look around [15,16]. This freedom of head rotation means the viewer now controls where to look in the scene, changing the viewer's perception of their role, extending their interactivity and agency [17]. Secondly, a CVR viewer is not a fully active participant, because CVR is a "lean-back medium" with limited interaction possibilities [18]. The viewer is still on the passive side and mainly wants to watch the story unfold and follow the narrative instead of acting in it [10]. We also found very little discussion in the literature that focused on the properties of interaction design from the perspective of user experience (such as the viewer's awareness of the affordances for interaction and expectations of agency). Discussion has typically been around how to enable full user interactivity in immersive environments [19,20]. However, full interactivity would require complex hardware and an overly-demanding amount of effort on the part of the viewer, and thus is not applicable to CVR. This amount of interactivity borders on something other than cinema, and reaches more towards an immersive video game. We consider CVR and video-game "users" to mainly have different motivations, so they likely comprise different demographics.

The remainder of this paper is organized as follows: We first review the literature covering CVR user experience and system design, immersive storytelling, traditional filmmaking, and game design. We then explore the continuum of viewer interactivity in CVR, summarizing the methods and studies from the literature. We then propose a Framework of Mediated CVR to create a CVR experience with implicit interactivity, aiming to balance the viewer interaction and agency, against the director's authorial control that is also necessary for storytelling, resolving the issue of the Narrative Paradox. The conclusion and future work are introduced in the last section.

2. Methodology

For this review, we mainly look at previous work related to CVR, from the perspective of narrative design and user experience design, rather than technical optimization or system building. We therefore did two rounds of search in the Scopus bibliographic database (<https://www.scopus.com/search/form.uri> accessed 18 May 2021), using the following combination of keywords. In round one, we used the search terms "360-degree" AND "storytelling". In round two, we used "cinematic virtual reality" AND "interaction". We

searched in title, abstract, and keyword fields among papers published from the year 2000 up to the present.

In the first round, we mainly looked for papers covering 360-degree video content creation, and obtained 37 initial results. We then used the sorting tool from the database to filter and keep only those that have been cited five times or more, resulting in 14 such papers left in this category.

In the second round, we extended the coverage of media types from only 360-degree video to any other immersive content with a narrative purpose. We also paid special attention to those which also involved “interaction” as their focus. The initial search returned 58 hits. Considering this is a relatively new research area, especially with the term “CVR” being brought forward around 2015, we did not filter them with citation numbers. Instead, we did a further review of those papers and removed several special cases that fell outside the scope of this research. These included (1) those which initially used an immersive environment for capturing, but rendered only a part of it as a 2D viewport for viewing by end users, (2) those which focused on technical optimization rather than user experience, and (3) those which worked with sound instead of visual presentation. This refinement resulted in 31 papers.

Finally, we checked for duplications from these two rounds, as some of them covered storytelling with both 360-degree videos and viewer interaction. This resulted in a final set of 41 papers that we reviewed in detail. This final set is also listed in Table 1, preliminarily grouped by the type of media, design focus and viewer interaction techniques. In the following chapters, we first review those around user experience and narrative design of storytelling with 360-degree videos. Then we move on to reviewing a series of works of narrative VR experiences, including but not limited to 360-degree videos, embedded with different levels and various types of viewer interaction techniques.

Table 1. Summary of the papers we review by types of media, their design focuses and viewer interaction techniques.

Media Types	Design Consideration	Viewer Interaction	Interaction Techniques	Papers in This Category
360-degree video	Viewer’s role	No	None	[2,9,10,15–17,21]
360-degree video	Placement of story elements (<i>Mise-en-scene</i>)	No	None	[7–14,22–24]
360-degree video	Viewer interaction	Yes	Pointing and activating	[2,25–29]
Synthetic 3D environment	Narrative with interaction	Yes	Head, eye, gaze, pointer	[18,30–34]
Synthetic 3D environment	Narrative with interaction	Yes	Text input, conversation, drawing	[35–41]
360-degree video or synthetic 3D environment	User input	Yes	Gestures, annotations	[19,20]

3. Moving from the Flat Screen to an Immersive Medium

In this section, we mainly look at previous work focusing on the CVR user experience and system design. Researchers have been migrating filmmaking grammars, principles about elements that contribute to the setup of the viewer’s experience, and the perception of affordance to interact into CVR, all aimed at relieving the Narrative Paradox and guiding the viewer along the storyline.

3.1. Choosing the Viewer's Role

When moving from a 2D flat video to an immersive medium, not surprisingly, researchers have noticed that the first obvious change is the Point of View (POV). In an immersive medium, the viewer sits in the center of the scene, instead of looking at a rectangular flat screen. Syrett et al. [17] stated that with this new POV, the viewer becomes the narrator, since she can choose what to look at and what to understand. This represents a change in the viewer's role. One of the commonly used methods to cope with this change is to define the viewer as either a spectator (not part of the story) or a character (either invisible or acknowledged) in the story. Bender et al. [21] compared the effects of two camera positions in CVR, the character view and the immersive passive view. They looked at their effects on attention to important spots. The character's first-person view did help the viewer to establish a fixation on Regions of Interest (ROI) faster than a third-person view.

Dooly et al. [9] in their research also proposed a division that is less coarse. They categorized the roles of a character into a silent witness, a participant, and a protagonist. The difference between a participant and a protagonist is whether other characters in the story give the viewer social acknowledgment, and how often and with what priority this is given. Furthermore, acknowledging the viewer as a person or a character is a CVR approach borrowed from theater practitioners. In staging and performing 360-degree video using principles from theatre, Pope et al. [7] found that actors had lines directly spoken to the viewer/camera, treated it as another person in the scene, and also changed them into an invisible one or a spirit, since in 360-degree video, the viewer cannot see her own body when looking down. Brewster [42] also stated that the viewers can alternatively take a "Morphing Identity," where they find themselves being something else in the story or the scene, rather than a human being. One example is the 360-degree film *Miyubi* (<https://www.oculus.com/experiences/gear-vr/1307176355972455/> accessed on 24 October 2018) from Oculus studio. In the film, the viewer finds that she has become a toy robot newly bought by a kid, and witnesses how other family members treat it. Figure 1 shows a screenshot from the film, with the toy's robot arms visible in the lower part.



Figure 1. A screenshot from the 360-degree film *Miyubi* (source: [Oculus.com](https://www.oculus.com/experiences/gear-vr/1307176355972455/)). The viewer is embodied as a toy robot. She can see the robot arms if she looks down. They are also visible in the screenshot. The image is slightly distorted because of the conversion from 360-degree video to a flat image.

Other than defining the viewer's role from a screenwriting or theater staging perspective, the pose a viewer takes to watch the narrative content will also influence how her role is perceived, thus affecting her behavior when watching. Godde et al. [10] observed viewer behavior under the same content, but in a seated pose vs. a standing pose. They point out that when seated (the most common way to view CVR), viewers spent a larger amount of time looking towards the front and showed less exploratory behavior. Tong

et al. [15] also explored the preferred user scenario for 360-degree videos, and coined the term “Swivel-chair VR” to specify the preferred way to consume 360-degree video content (seated, instead of standing). Swivel-chair VR describes a scenario where the viewer watches a 360-degree video while sitting in a swivel chair, wearing a VR headset. The swivel chair allows the viewer to rotate around 360 degrees by turning the body together with the chair. It is easier and more comfortable than only turning with one’s neck, as the swivel chair serves both as a cue to imply the affordance of rotation, and an anchor point to assist rotation. It also provides greater comfort and safety, compared to standing while viewing.

3.2. Choosing the Placement of the Camera, the Actor, and Other Story Elements

Mise-en-scene has been a powerful tool for filmmakers to design story elements to help the viewer find out where she is located and what to focus on in the scene. This was also explored in the context of CVR. Dividing zones around the viewer was proposed because viewers were given the freedom to look around and decide where the current view and focus would be in CVR. However, researchers discovered that not all directions around a viewer had equal weight, given a choice. Godde et al. [10] divided the full 360 area around the viewer into the front area, rear area, and a blind spot directly behind, as shown in Figure 2. They point out that viewers mainly tend to focus only on elements within the front zone (180-degree front-facing) and are less likely to look for elements inside the rear zone, where significant head turning is needed. The blind spot is where elements will most likely be ignored or missed by the viewer, even though they might be related to the narrative. With proper use of staging and directing cues, viewers can be encouraged to explore the rear zone, but the placement of important elements in the front zone is still recommended, as supported by their experiment. They also found that if the viewer takes a seated pose, the preference for the front will be intensified.

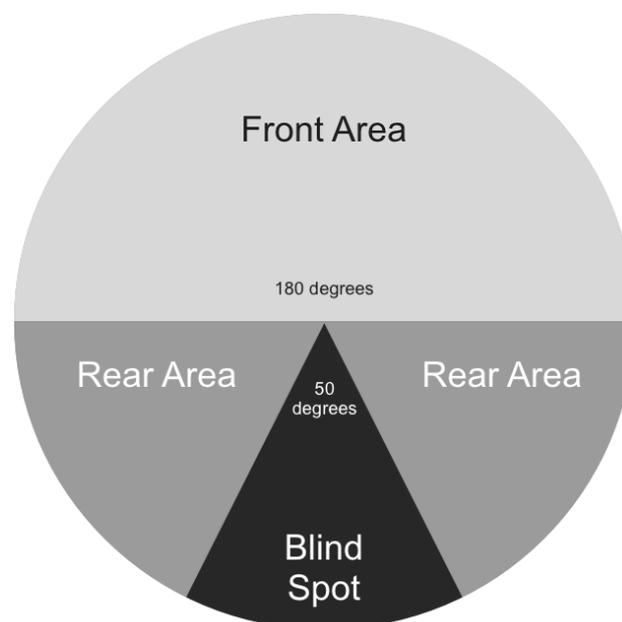


Figure 2. The division of staging zones in the full 360 area around the viewer by Godde et al. [10]. Notice, there is a blindspot directly behind the viewer, of about 50 degrees.

Distance is another key factor to be taken into consideration. Dooley [9] pointed out that in traditional filmmaking, within the border of a frame, directors can use “shots” to control what story element to focus on, such as a close-up or a wide view, to direct the viewer’s gaze and attention. Likewise, in CVR, while there is no frame to define a shot, defining “experience” with “distance between the viewers and actors” can be important and viable. They proposed the theory that “in CVR distances can influence the viewer’s

emotional engagement with the characters” and analyzed three scenes in a sample narrative 360-degree video *Dinner Party* (<https://www.with.in/watch/dinner-party> accessed on 1 March 2021) to see if distance variations changed how viewers perceive their own relationships with the characters around them. In the experiment carried out by Pope et al. [7], they asked actors to stage and perform a short drama for either a viewer sitting on a swivel chair or a 360-degree camera taking the place of the viewer. They discovered that when the actual viewer was replaced by the camera, actors still performed with the principles regularly used in a theatre. The actors tend to group on one side of the camera so the viewer could easily see all the action without turning her head frequently. In another test, Bailenson et al. [22] also discovered that in an immersive environment, users tend to maintain personal space and keep their distance from other human characters just like in real life. They also discovered that if the other human characters are making eye contact with the user, this tendency of distance-keeping will be more obvious. However, unlike CVR, their work was mainly conducted in a highly interactive setting where the user could freely move around.

Other than controlling from the *Mise-en-scene* perspective, directors can also manipulate the parameters of the camera directly, to change how the viewer perceives her role. The height of the camera was the first factor to be addressed. Researchers have discovered that the height differences between the camera and eyes are more accepted by the viewers if the camera position is lower than the viewer’s eye height [8]. Additionally, a seated pose is preferred and can be adapted to more easily than a standing pose [23].

4. The Continuum of Interactivity of Immersive Experiences

As mentioned in the previous section, once moved from a flat screen to an immersive medium such as the 360-degree video, the viewer’s role changes. The shift means the viewer is now part of the story world (as a character or not) and will expect more agency [1,30]. We can infer that the viewer will want to interact with elements in the story world and influence the narrative. Thus, viewer interaction in CVR is another dimension we need to look at. Viewer interaction, in the setting of immersive storytelling, stands for a certain control the viewer has over the narrative [35]. This control differs from one system to another (either because of its hardware or the purpose of storytelling). In the following sections, we would like to first propose the idea of a “Continuum of Interactivity” to divide and group the experiences and projects we reviewed by the level of interactivity viewers have in each of them, from Very Limited to Limited, Medium, and High, by placing them on the continuum, as shown in Figure 3. We will then discuss each group by looking into their interaction design decisions and properties.

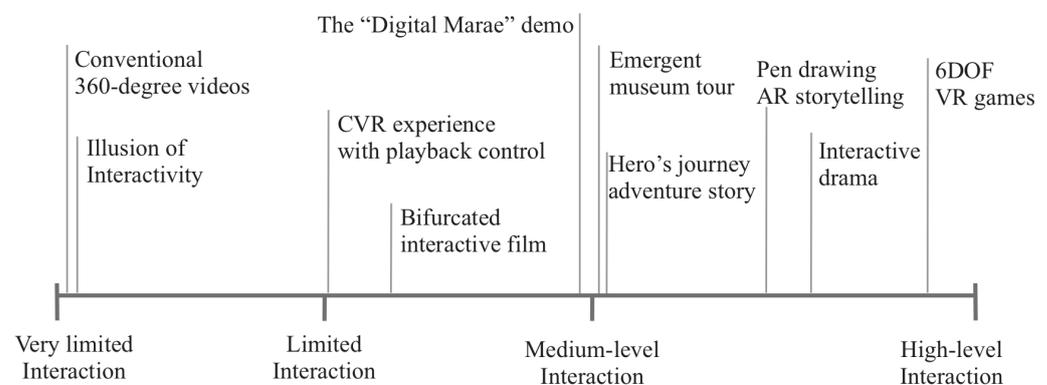


Figure 3. The continuum of interactivity. On the horizontal axis from left to right, we put a variety of immersive experiences with different levels of interactivity, not limited to only CVR, but also VR games, interactive theatre, etc. A spot on the horizontal axis represents the level of interactivity an experience has, varying from very limited to highly interactive.

4.1. Zero Interactivity

The most common and easily accessible form of CVR is 360-degree video. Like traditional filmmaking, the viewer takes a seated or standing pose, watching a 360-degree video with either a flat-screen device or wearing an HMD [15]. In both cases, the only user input is to choose in which direction to look at any given time. In a traditional 360-degree video, this input will not affect any narrative progress, and the director does not predefine any reactions to it [13,24]. One can conclude that the viewer in a 360-degree video playback has zero interactivity with the narrative. Those experiences are placed on the left-most part on the continuum as shown in Figure 3.

4.2. Illusion of Interactivity

Other than addressing the directors and actors, we also found that some researchers try to embrace the viewer's first-person view using an "illusion of interactivity" approach, which helps to increase the level of spatial presence and realness, even when no "real" interaction is taking place (cf. [43]). Brewster from Baobab studios [42] presented the method they applied in the computer-generated narrative VR short film *Asteroids* (<https://www.baobabstudios.com/asteroids> accessed on 20 April 2018), where a robotic dog mirrors the viewer's simple head tilting during a dwelling stage before the main story starts. The mirror action from the dog gives the viewer a further impression that she is part of the scene and feels like her movement can affect the scene itself, whereas general interactivity is actually not enabled and the experience is still a linear film. Dooley [3] also point out that a strategic VR director can create the illusion of choice for the viewer, when in fact they are creating a series of audio and visual cues that result in a preconceived narrative experience. This can be done by deliberately laying down a series of content chunks with auxiliary transitions and assisting content to fill the blanks between important story nodes. Thus, the viewer will feel that she is in control and that the story unfolds because she noticed something special or made the choice of looking at a certain object first. However, system wise, the viewer is not actually interacting with the environment, nor do they have any impact over how the narrative unfolds. Thus, they are also grouped to the left-most part on the continuum.

4.3. Medium-Level Interactivity in CVR

Acknowledging that viewers intend to interact, researchers have been exploring and evaluating various interactive techniques for CVR. Depending on the content, the genre, or the director's specific intention, each approach registers its own combination of choices on several properties. We focus on two of them here, the visibility of the interactive element and the depth of impact on the narrative.

Before jumping into design choices for interaction in CVR, however, we first look at the tasks CVR viewers carry out and the input they employ to perform them. In a technical survey, Roth et al. [18] point out that the main interactions in CVR, if enabled, are: selecting visible images, selecting areas for nonlinear stories, triggering the next scene or displaying more information, and navigating in the movie (menu). These were commonly implemented in, for example, "interactive 360-degree videos" or similar immersive experiences with content either pre-recorded or computer generated. They also listed two types of input based on the tasks above, *continuous* input, mainly for tracking, moving (objects or as navigation tasks) or pointing, and *discrete* input, which is for trigger actions and activation. To carry out these inputs, a CVR viewer may use head rotation, eye gaze, hand gestures, or a pointing device, depending on the system design.

Generally, when creating a CVR experience, the director will need to go through three design decisions before finally implementing the specific interaction technique: the *tasks* she wants the viewers to carry out, the *consequence/impact* each task will have, and the *input method* the viewers will use to perform them. In the following sections, we look at several examples with various interaction designs. We also pay extra attention to their visibility to the viewer and their impact on the narrative itself.

4.3.1. Shallow Intrusion—Temporal Control

A typical user input in CVR is through temporal control, which means the viewer can only interact with the playback of a narrative (such as speed change, pause, and browsing), but is unable to break the linear progress. In a conventional 360-degree video player, a viewer is presented with a “menu bar” to interact with, as shown in the two examples from Pakkaen et al. and Keijzer [25,26]. Researchers developed and tested various input techniques, but the control panel stayed similar. They were always derived from the familiar desktop UIs. One requirement of these UIs is that eye and hand coordination is needed to complete the “pointing-to-activate” two-step action, temporally taking away the viewer’s capability of making spatial choices and looking at the scene itself. Petry et al. [27] invented a new system that decoupled orientation control from temporal control. They kept free head rotation to control where to look and at the same time enabled an extra pointing gesture for fast-forward/rewind. The viewers were therefore given parallel capabilities to look around and browse through events chronologically at the same time. These two inputs also did not interfere with each other, unlike the examples given previously. However, all of them only brought the viewer to the level of controlling the “temporal progress” along a linear track; the storytelling itself was pre-recorded and fixed. Thus, the interaction was “shallow” and “limited.”

4.3.2. Deeper Intrusion—Narrative Control

Creators are also aware that, in CVR, after acknowledging the viewer as a character in the scene, adding interaction enhances the viewer’s feeling of presence, because having an active role contributes to their enjoyment and engagement [28]. At the current stage, as most CVR content is pre-recorded, the possibilities of interaction with the narrative itself are mainly limited to two: (1) choices over a bifurcated plot where every scene is a video clip, and (2) the overlapping of extra elements over each video clip, injected into the scene [30]. Interactive narrative content, also known as Interactive Fiction [31], is a form of narrative based on a bifurcated story and has been commonly available. Content can be found in both traditional flat-screen media, such as the sci-fi drama *Bandersnatch* from the series *Black Mirror* (<https://www.netflix.com/sg/title/80988062> accessed on 6 May 2019), and in immersive media, such as the virtual relic city tour *Bagan* (<https://artsexperiments.withgoogle.com/bagan> accessed on 14 January 2019) based on recorded 360-degree videos and rendered 3D scenes. In interactive experiences, the viewers make choices at each “intersection” (refers to story nodes in the design) and rearranges the linkage of fragments into their own configuration [32]. One example is the previously-mentioned drama *Bandersnatch*, in which the viewer will occasionally be presented with two choices throughout the story, as shown in Figure 4. Each interaction inside the experience is reactive, from a technological point of view. However, this is challenging from a narratological/authorial point of view, in terms of keeping the story flow and maintaining user engagement with the narrative, because the creator of the story cannot predict each viewer’s actual chain of choices when the content is being presented. Thus, the authorial control over the narrative is disrupted.

4.3.3. Explicit and Implicit Interaction

In typical designs of viewer’s interaction with the user interfaces of the narrative system, visible elements, such as a circle target, a countdown marker, or a translucent dot [33] assist the viewer in becoming aware of the location of an ROI, the effectiveness of a recent input, or a commencing event. When novel designs started to move away from conventional UI, people also explored the possibility of enabling user input without relying on any visible interface or reaction element. Roth et al. [29] list parameters around the activation targets. Targets can turn visible when triggered by the user, or remain invisible, depending on the requirement of the narrative itself. In most narrative VR, like with movies, even a small non-diegetic object can be disturbing and break the feeling of

presence. Thus, the trigger or target will need to be only visible when activated or invisible throughout the entire experience.

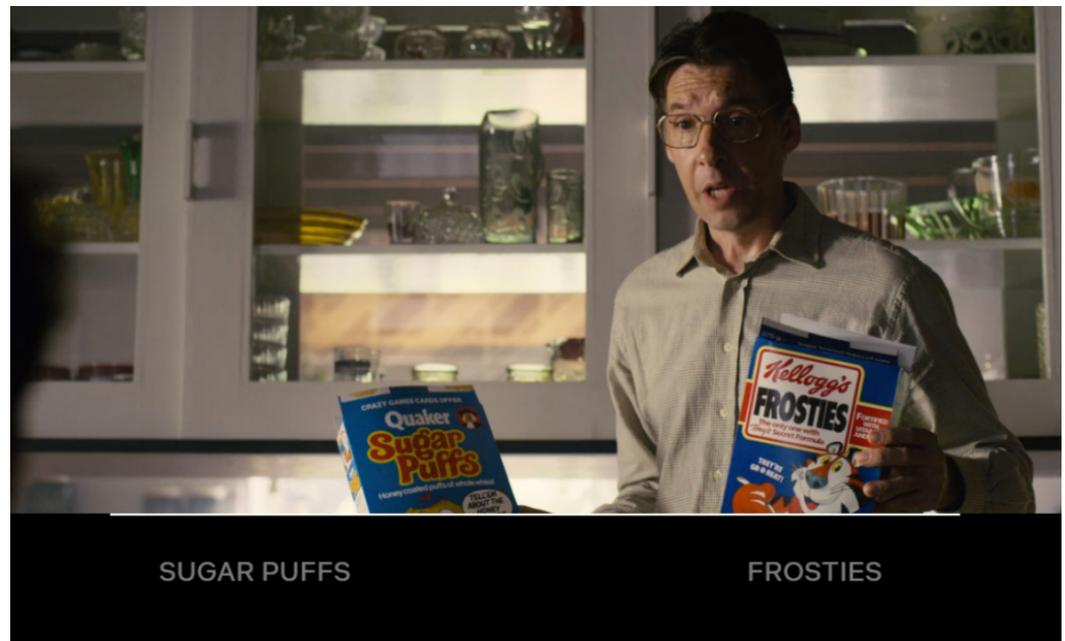


Figure 4. Screenshot from the interactive film *Bandersnatch* (source: Netflix). At several given moments in the film, the viewer is presented with two choices. It will impact the current character's decision of imminent action (or may not). A countdown timer is also presented (the thin line above the options). If the viewer does not make a choice before the time runs out, a default option will be chosen automatically.

In an experiment conducted by Ibanez et al. [34], they constructed a virtual tour system that generates stories based on the location designated by the user and the location where a virtual tour guide was standing. Therefore in the entire tour, the user feels like she is guided by a real tour guide, with all the knowledge and responses to her choices of POIs instead of a linear pre-recorded video footage. In fact, however, all the granular narratives are pre-fabricated. Therefore, compared to a conventional system where the viewer (user) only gets to navigate along a one-dimensional timeline with explicit input, in this virtual tour system, the viewer naturally browses the scene, and the narrative structure changes accordingly on the fly. During the entire process, the viewer is providing implicit input to the system (naturally choosing where to look and focus, as one would do in a real-world tour), and is unaware of the fact that she is making choices. Another example of this implicit interaction is a six-degree-of-freedom (6DOF) "Digital Marae" experience that we have implemented (cf. [44]). A *marae* in this context refers to physical Māori location containing a complex of buildings around a courtyard where formal greetings and discussions take place. A voxelized avatar of a real person from the marae in the physical world acts as the host in a computer-generated environment of the marae. The host delivers narratives while the 6DOF viewer is freely roaming in the scene. Three-dimensional clips of a real storyteller (*voxelvideos*) are rendered (visually and acoustically) in the virtual marae environment. The virtual storyteller introduces artifacts and decorations in that marae, as shown in Figure 5. This can be stopped and started with explicit user interaction, similar to the UI controls described in the previous section. As an option, the voxelvideo host avatar is also capable of actively establishing eye contact with the viewer and initiating the introduction when she is within a certain proximity, maintaining eye contact when speaking to the viewer. When the viewer steps away from the gaze-maintaining proximity, the host avatar goes back to the initial pose (of not attending the viewer) and pauses the speech. This proximity setup gives the viewer the impression that she is interacting with a real person-like host

when she roams the marae, and that the host is giving a tour especially for her, instead of the impression of “a holographic recording that is played when I press a button”.

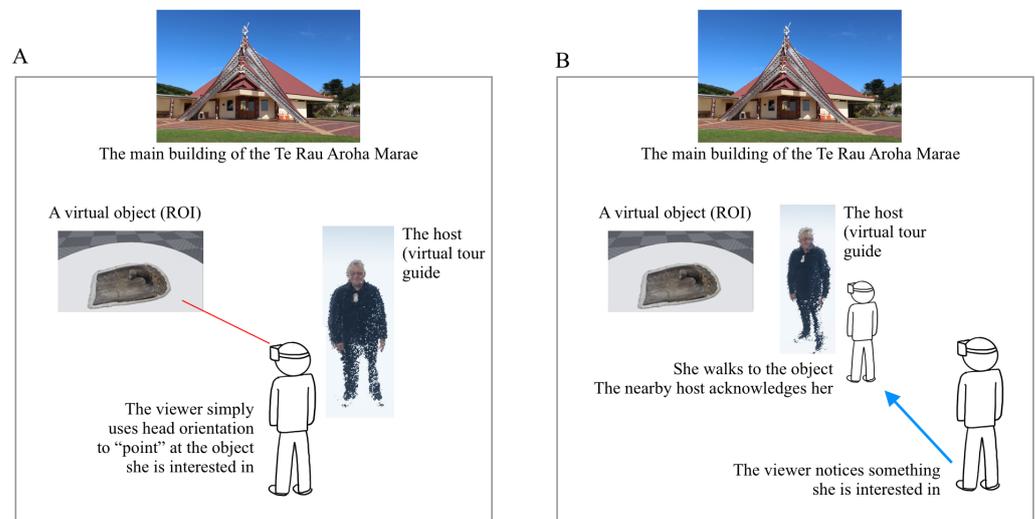


Figure 5. Two illustrations of the Digital Marae experience, of the Te Rau Aroha marae (source: Park et al. [44]). In **setup A** on the left, the viewer uses head orientation to “point” to one of the objects (ROIs) she is interested in, in this case, a ta tā or tīrehu (bailer). The virtual host (voxelized avatar) stands beside the viewer to deliver the narrative. In **setup B** on the right, when the viewer notices an object she is interested in, she moves (or teleports) to the proximity of that object. The virtual host stands next to the object, acknowledges the viewer when she is nearby, and delivers the narrative.

In these two examples, although the viewers are not aware of any interactive elements, nor any trigger activation that is revealed to them, they are indeed interacting with the system and affecting the process of storytelling; thus, both are placed near “medium-level interaction” on the continuum (Figure 3).

4.4. High Level Interaction

Simply browsing pre-recorded clips by pointing and clicking is not the only type of immersive storytelling. On the opposite end of the continuum, researchers have also tried advancing narrative experiences by employing interaction techniques of high-level fidelity. Examples include emergent storytelling, interactive drama, and sandbox video games. Sharaha and Dweik [35] reviewed several interactive storytelling approaches with different input systems. They have one aspect in common, namely that the users are highly interactive since they act as one of the virtual characters in the story. In the system designed by Cavvaza et al. [36], they focused on auto dialogue generation to drive the narrative based on the text input provided by the user. A similar approach can also be found in the well-known video game *Façade* [37,38], in which the story unfolds as the player interacts with two virtual characters by typing texts into the console (sample screenshots of the game are shown in Figure 6). Edirlei [39] proposed another storytelling experience where the viewer watched the narrative via an Augmented Reality (AR) projection system and participated in it by drawing. The virtual characters were performing, and the viewer could physically draw objects on a tracked paper and then “transfer” it into the story scene as a virtual object to interact with the characters and the storyline. In another large-scale CAVE-like experience [40], the story unfolded as the user was having natural conversations with another virtual character. The user was able to freely speak and move around in the scene. The system reacted to her behavior and progressed with the narrative. In those experiences, the viewers (or users) were given interactivity of high fidelity (speech, gestures, locomotion, drawing). Because of the abundance of input and interaction techniques a player (no more simply a viewer) could use, these experiences are placed to the right end of the continuum, as shown in Figure 3.



Figure 6. Screenshots from the video game *Façade* [37] showing the two characters. The player interacts with them by typing dialogues into the game window. The conversation is visible on the screen.

5. Discussions

In the previous section, we stated that the preferred user scenario for CVR is the “Swivel-chair VR”, where a passive viewer sits on a swivel-chair (or, less likely, standing). The viewer expects some agency from the storytelling but still wants to enjoy the narrative at a “lean in” mindset instead of “lean forward” [45]. If we look at the continuum (Figure 3) and consider where “CVR with interactivity” should sit, we can see that “very limited interactivity” is not the desired place to implement narrative VR. In many CVR experiences, the level of interactivity provided by the system does not match the viewer’s expectation of participation when moving from a flat screen into this immersive medium. “High-level interactivity” is also not applicable because in CVR, especially the preferred “Swivel-chair VR”, the viewer’s capability of interaction, objectively, is still insufficient to perform full-body action or delicate operation on objects. Furthermore, subjectively, the viewer is still passive, unwilling to act with full interaction, but rather to enjoy the story unfolding in front of her. We can therefore preliminarily infer that a proper system design for “CVR with interactivity” will be placed at either the spot near “limited” or “medium” on the continuum.

A storytelling framework presented by Reyes [30] fits well into this category. She presented the possible diegetic interaction options on a pre-scripted story with different navigation alternatives. The aim is to produce an interactive narrative that is independent of the user’s journey within the story; the plot is always created with a dramatic climax, or say, “ensure[s] the linear progression of the dramatic arc independently of the journey shaped by the user’s choices”. In her proposed system, she put forward this idea of “limited interactivity”, as the user has some level of free choice, but the general narrative structure is based on the “hero’s journal” and is controlled and made by the director (like the double-diamond shape approach of in video games [46]). The primary nodes are always defined (and will be a key element to drive the story forward), with secondary nodes of different paths in between for free choice. In their paths, two types of links were designed and implemented, where “external links” are those jumps between pieces of the stories, moving alongside the general story arc, and “internal links” are extensions within a node, pointing to extra pieces of stories but not critical to the arc (providing more information to enrich the experience). A similar approach has been evaluated by Winters et al. [4]. They used structural features of the buildings and terrains in a 3D game world to implicitly guide the players along a path that was preferred by the storyteller. They also placed landmarks with outstanding salience in the backdrop of the game world so that, while the player can freely roam the world, she was still attracted by the landmark and would eventually go to the key place where the main plot is taking place to drive the narrative forward.

When reviewing the studies in the literature and grouping them by looking at the “level” of interaction a viewer or player has during the experience, we also noticed that a single-dimension continuum might not be the absolute metric to use when exploring viewer interaction in CVR. As we stated, on one hand, we saw that the forms of viewer interaction varied among very limited to medium-level to high and abundant, determined

by system design; on the other hand, similar forms of interaction techniques also have a different impact on the story itself, such as simply scrubbing around the timeline [25] or affecting how the story unfolds and ultimately changing the outcome, such as in *Bandersnatch*. The “level of interactivity” or say the “complexity of the interaction technique” seems not strictly equivalent to the “depth of impact”, as we described in a previous section. It is possible that when an interactive storytelling system is implemented, its input methods and interaction techniques define its position on the “continuum of interactivity”. However one can not fully predict how the viewer will choose to participate and impact the narrative, because the actual use case varies from one viewer to another. We have seen discussion from researchers who looked at both technical and narrative aspects of immersive experiences separately [47]. Koenitz [41] also put forward a theory indicating that the viewer’s participation in storytelling is also part of the narrative itself. We assume further exploration will be needed on the relationship between the viewer interaction design choices and the actual viewer experience.

Furthermore, from the aforementioned interaction design case, we had this observation that there is no definite “better” interaction design for satisfying user experience. In one of our ongoing research projects, one iteration of the Digital Marae experience (cf. [44]), we administered two setups: (1) swivel-chair CVR where the viewer used head orientation for “pointing”, (2) walk-around CVR where the viewer was given a controller for “point to teleport”. The viewer’s capabilities were also different in those two setups. In the first setup, the viewer’s input directly indicated the content she was interested in at a given moment. The viewer had definite control over the sequence in which the stories of ROIs were presented, but there was no priority between one against another. Their spatial relationship was also not preserved. The viewer felt she was browsing a “flat image” of the entire marae. The guidance of the host was diluted, thus leading to lower narrative engagement. In the second setup, the viewer was able to spatially get “closer” with one of the ROIs and “further” from others. The system interpreted the viewer’s choice by measuring which ROI she was closest to and presented the story related to that one. Compared to the first setup, the viewer had a higher level of narrative immersion because she felt the host knew where she was and what her focus was as she roamed the virtual environment. However, this significantly increased the complexity of the system as the viewer needed input for teleportation and the director needed to cope with extra factors such as the viewer’s distance and possible interruptions when the story about one specific ROI was being delivered. We think that there is no “correct” choice between those two setups, but it all depends on the director’s choice of the purpose of the installation and what experience the viewers were expected to have.

Therefore, at this stage, we want to preliminarily conclude the following. (1) A CVR viewer is a “lean-in” viewer. It is recommended to consider the viewer as a character in the scene who has a certain level of impact on the narrative, which generates user agency, presence, and engagement. At the same time, the viewer’s participation needs to be considered because her role is different from the one in cinemas or video games; (2) to enable interaction, the creator will need to consider the tasks a viewer wants to carry out by interacting, the impact over narrative each of those tasks will have, and the input method the viewer will use to perform those tasks. Especially on the input method, we have seen both explicit and implicit ones. Based on conclusion 1, since the viewer is only “lean-in” instead of fully willing to participate, implicit techniques are recommended; (3) authorial control is still necessary for the delivery of a complete story. The director still needs it to construct the story arc to evoke emotional engagement in the viewers. Internal links between the story elements need to remain intact despite the viewer’s interactivity; (4) whether a higher level of “abundance and complexity of viewer interaction” will necessarily lead to a higher level of narrative engagement and enjoyment remains unclear. Directors need to consider this under the general purpose of the storytelling activity itself.

6. Adaptive Playback Control as a Concept Framework

In this section, we propose the concept of Adaptive Playback Control (APC) as a machine-mediated narrative framework. We put it forward as a conceptualized solution of enabling viewer's interaction in CVR and as an example case to those concepts we stated, especially a representation of the exploration of the fourth item in the preliminary conclusions from the previous section. This narrative framework is put forward with these considerations: (1) the viewer is using a (seated) swivel-chair VR experience with an HMD; (2) the viewer is passively watching, but the screenwriting from the director will define the viewer as (at least) a participatory level character in the story; (3) the viewer's experience will not be interrupted by non-diegetic elements, and the time flow will not be explicitly broken to solely "compensate" interaction.

We looked at both ends of the continuum we proposed, considered the viewer's expectation of agency and the director's expectation of authorial control, leading to three system components:

- A *story builder* component for the director to fill in raw materials and references of the internal relationships between key story nodes and their weights of priority;
- A *moderator* component runs in backstage to capture the viewer's real-time behavior and mediates the playback process as responses to the viewer's actions;
- A *projector* component presents the content to the viewer, and also receives commands from the *moderator* to update the presentation when necessary.

The director provides the actual content in the form of nodes, plus "map-like" references for the *moderator*. The "map" defines remarks on the nodes and paths with "weights of necessity" such as which of the nodes are mandatory and what parts of the path are pre-defined and cannot be waived. The *projector* captures implicit inputs from the viewer (such as head rotation and gaze change), hands it over to the *moderator* to be read as signals, and alters the path of how the story progresses from one key node to another in real time. Figure 7 illustrates the components in this system and how they work together to deliver the final result to the viewer. We came up with the term "Adaptive Playback Control (APC)" to describe the entire structure.

One of the plans is to integrate this narrative system into the Digital Marae experience we introduced in the previous chapter. On top of the eye-contact setup, we plan to add another layer to the viewer's capability of freely roaming the marae. Pre-recorded voxelized clips of the host introducing a list of ROIs (such as ritual objects, decorated walls, sculptures) and the host performing "functional events" such as asking or waiting for a certain action, as well as idle state, are all prepared and stored in the *story builder*. When the viewer is walking and looking around in the Digital Marae, the *projector* monitors the real-time location of the viewer and triggers an "inquiry event" when the viewer approaches one of the feature points and ponders for a certain period of time. The avatar of the host will approach the viewer and ask if she wants to know more about that feature point, just like one will experience in a museum when a staff member is offering help (on the background, the system is playing one of the "functional event" clips and waiting for the viewer's response). If the viewer confirms with certain input (head nod, or simple gesture, depending on the system configuration), the *moderator* will make decisions and branches out to play the clip in which the host will start to tell the story related to that object in front of the viewer. To make the experience more like real-life events, the *moderator* will also pause the narrative when the *projector* detects the viewer is looking away and "appears to be no longer interested in the current content". Generally, the experience is played out around the viewer's behavior and actions in the scene. The three components of the APC system work together to shuffle and reconnect the clips provided by the creator into one continuous tour-like experience.

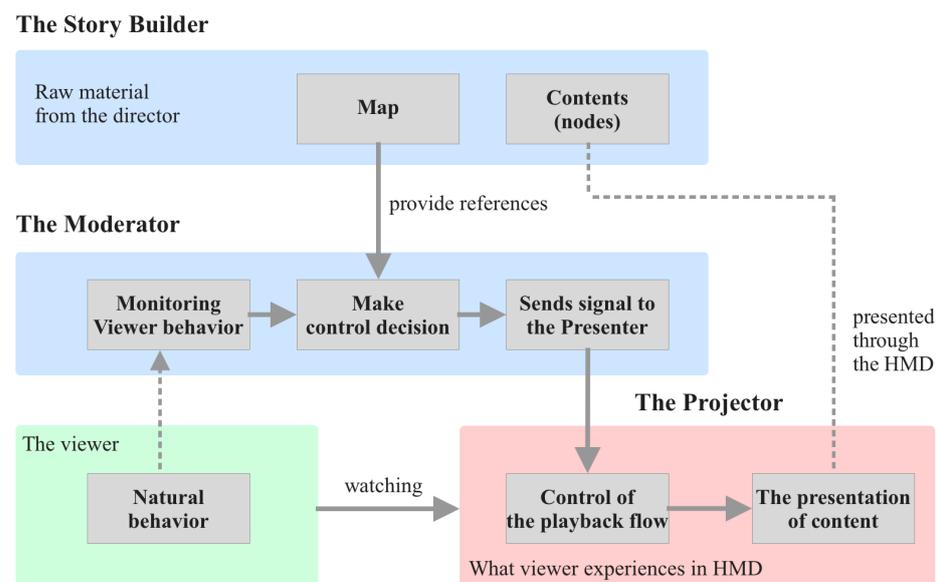


Figure 7. The structure diagram of the APC system. It shows the three components in the system, a viewer who is using the system, and how the components work together to deliver the final result to the viewer. From top to bottom: Before the actual playback, the director feeds the contents (as nodes), and a “map” reference to the *story builder*, as raw materials. The “map” is a reference for the *Moderator*, with remarks of possible paths a viewer can later take connecting the nodes in an actual playback, the mandatory paths, as well as their priorities. When an actual playback starts, the viewer watches the content, wearing an HMD. At the same time, the *moderator* monitors the viewer’s behavior and analyzes which node the viewer is interested in, cross-references with the “map”, and determines which one of the nodes will be presented next. Then, the *moderator* sends the directives to the *projector*. The *projector* responds to it and presents the chosen node (content) to the viewer. Since everything is running in backstage, the viewer is unaware of the fact that she does make choices, albeit implicitly.

In an immersive experience like this, the viewer can find herself taking a role in the story, with a certain level of interaction with the system, and finds agency with it, because the path of how the story progresses forward does change (within certain boundaries) according to the viewer’s behavior. On the other hand, the main purpose of the storytelling and the emotional response the director wants to invoke in the viewers can still be ensured because the essential narrative arc and the main plot were still laid by the director and not affected by the viewer’s actual behavior when watching. We can also observe whether the viewer’s feel of narrative immersion and feel of engagement and enjoyment will be influenced by the fact that the interaction she has within the story world is at a “limited level”, compared to the 6DOF version we presented previously.

With this framework, we want to “interactivize” the CVR experience for the intrinsic nature of immersive media, such as 360-degree video and computer-generated immersive film, but not for the full interactive experience. This is because high-level interaction requires complex elements such as full body avatar, emergent storytelling mechanisms, AI components and other advanced technologies, which are not supported by most CVR experiences. This “one step back” is also due to the prospective application we want to explore. We aim to provide a framework for immersive and interactive content creators who develop engaging and enjoyable experiences for entertainment, learning, or invoking empathy. These include teachers who want to create a visual demo for their classes or museum curators who want to create virtual tours for online and remote visitors. We expect that such a framework will give creators a familiarity akin to scripting for conventional videos. Thus, on one hand, it is still a pre-scripted narrative at the backbone but an interactive and immersive experience at the front face. It can ensure the narrative arc remains in control of the director, but the freedom of interaction is in the hands of

the viewers. We expect that this is a solution to archive the balance and relieve the Narrative Paradox.

7. Summary and Future Directions

In this paper, we reviewed literature about designing narrative VR content with the consideration of the viewer's role and viewer's interactivity, aimed to resolve the issue of the Narrative Paradox. We first looked at previous work by grouping them under the *Mise-en-scene* and *cinematography* of traditional filmmaking. Researchers have explored the viewer's role in CVR and its relationship to the characteristics of other story elements in the scene when interactions are limited. From their insights, we also learned that theater practices can provide a good reference to the configuration of placement and distance of essential story elements to clarify the viewer's role and the focus of the story.

Then, we moved on to review the literature around enabling viewer interaction in CVR. We proposed the "continuum of interactivity" of immersive experiences to place various types of approaches we visited, to categorize them by the "level of interactions" they have and see which is appropriate for creating CVR experiences with interactivity, to achieve high level engagement and presence. According to the continuum, we also discuss the "depth" of narrative impact those interaction designs have and their visibility to the viewer in the story world.

We made four preliminary conclusions after the literature review, covering the discussion around viewer's role in CVR, factors affecting interaction design, director's authorial control, and the relationship between interaction's level of complexity and its impact on narrative immersion. Then, the "Adaptive Playback Control (APC)" framework is proposed as a conceptualized example of enabling viewer interaction in immersive storytelling with backstage authorial control, plus consideration of viewer's role and the context of the story. This framework is yet under further exploration and tailoring. In the future, we plan to implement it in both 360-degree videos and computer generated 3D scenes. We will also conduct user studies to evaluate its effectiveness and gather user feedback. Overall, we aim to propose a framework of "mediated CVR storytelling" with "interactivity". We believe such a framework can contribute to mitigating the issue of the narrative paradox in CVR and can be used by content creators to develop engaging CVR experiences for education, entertainment, and other applications, without the need for professional knowledge in VR and immersive system design.

Author Contributions: Conceptualization, L.T., R.W.L. and H.R.; methodology, L.T. and R.W.L.; investigation, L.T.; visualization, L.T.; writing—original draft preparation, L.T.; writing—review and editing, L.T., R.W.L. and H.R.; supervision, R.W.L. and H.R.; project administration, H.R.; funding acquisition, R.W.L. and H.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the New Zealand Science for Technological Innovation National Science Challenge project Ātea.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank Noel Park and Stu Duncan from the University of Otago, Yuanjie Wu and Rory Clifford from HIT Lab NZ, University of Canterbury during the implementation of voxel-based telepresence system, and the support of other research members in the New Zealand Science for Technological Innovation National Science Challenge (NSC) project Ātea.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mateer, J. Directing for Cinematic Virtual Reality: How the traditional film director's craft applies to immersive environments and notions of presence. *J. Med. Pract.* **2017**, *18*, 14–25. [\[CrossRef\]](#)
2. Rothe, S.; Buschek, D.; Hußmann, H. Guidance in Cinematic Virtual Reality—Taxonomy, Research Status and Challenges. *Multimodal Technol. Interact.* **2019**, *3*, 19. [\[CrossRef\]](#)
3. Dooley, K. Storytelling with virtual reality in 360-degrees: A new screen grammar. *Stud. Australas. Cine.* **2017**. [\[CrossRef\]](#)
4. Winters, G.J.; Zhu, J. Guiding Players through Structural Composition Patterns in 3D Adventure Games. In Proceedings of the 40th International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH 2013, Anaheim, CA, USA, 21–25 July 2013. [\[CrossRef\]](#)
5. MacQuarrie, A.; Steed, A. Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. *Proc. IEEE Virtual Real.* **2017**, 45–54. [\[CrossRef\]](#)
6. Bordwell, D.; Thompson, K. *Film Art: An Introduction*, 10th ed.; McGraw-Hill: New York, NY, USA, 2013.
7. 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, Denver, CO, USA, 6–11 May 2017; ACM: Denver, CO, USA, 2017; pp. 4468–4478. [\[CrossRef\]](#)
8. 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, Tokyo, Japan, 28 November–1 December 2018; ACM: Tokyo, Japan, 2018; pp. 1–2. [\[CrossRef\]](#)
9. Dooley, K. A question of proximity: Exploring a new screen grammar for 360-degree cinematic virtual reality. *Media Pract. Educ.* **2019**, *21*, 81–96. [\[CrossRef\]](#)
10. Gödde, M.; Gabler, F.; Siegmund, D.; Braun, A. Cinematic Narration in VR—Rethinking Film Conventions for 360 Degrees. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2018.
11. Kjær, T.; Lillelund, C.B.; Moth-Poulsen, M.; Nilsson, N.C.; Nordahl, R.; Serafin, S. Can you cut it? An exploration of the effects of editing in cinematic virtual reality. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, 8–10 November 2017.
12. Fearghail, C.O.; Ozcinar, C.; Knorr, S.; Smolic, A. Director's cut—Analysis of VR film cuts for interactive storytelling. In Proceedings of the 2018 International Conference on 3D Immersion (IC3D), Brussels, Belgium, 5–6 December 2018.
13. Lin, Y.C.; Chang, Y.J.; Hu, H.N.; Cheng, H.T.; Huang, C.W.; Sun, M. Tell Me Where to Look: Investigating Ways for Assisting Focus in 360 °Video. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI '17, Denver, CO, USA, 6–11 May 2017.
14. Brown, A.; Sheikh, A.; Evans, M.; Watson, Z. Directing attention in 360-degree video. In Proceedings of the IBC 2016 Conference, Amsterdam, The Netherlands, 8–12 September 2016; p. 29. [\[CrossRef\]](#)
15. Tong, L.; Jung, S.; Li, R.C.; Lindeman, R.W.; Regenbrecht, H. Action Units: Exploring the Use of Directorial Cues for Effective Storytelling with Swivel-chair Virtual Reality. In Proceedings of the 32nd Australian Conference on Human-Computer Interaction, Sydney, NSW, Australia, 2–4 December 2020; ACM: Sydney, NSW, Australia, 2020; pp. 45–54. [\[CrossRef\]](#)
16. Gugenheimer, J.; Wolf, D.; Haas, G.; Krebs, S.; Rukzio, E. SwiVRChair: A Motorized Swivel Chair to Nudge Users' Orientation for 360 Degree Storytelling in Virtual Reality. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI '16, San Jose, CA, USA, 7–12 May 2016.
17. Syrett, H.; Calvi, L.; van Gisbergen, M. The Oculus Rift Film Experience: A Case Study on Understanding Films in a Head Mounted Display. In *Intelligent Technologies for Interactive Entertainment: 8th International Conference, INTETAIN 2016, Utrecht, The Netherlands, 28–30 June 2016, Revised Selected Papers*; Springer: Berlin, Germany, 2017. [\[CrossRef\]](#)
18. Rothe, S.; Pothmann, P.; Drewe, H.; Hussmann, H. Interaction techniques for cinematic virtual reality. In Proceedings of the 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019, Osaka, Japan, 23–27 March 2019; pp. 1733–1737. [\[CrossRef\]](#)
19. Teo, T.; Lee, G.A.; Billinghamurst, M.; Adcock, M. Hand gestures and visual annotation in live 360 panorama-based mixed reality remote collaboration. In Proceedings of the 30th Australian Conference on Computer-Human Interaction, Melbourne, VIC, Australia, 4–7 December 2018; ACM International Conference Proceeding Series.
20. Piumsomboon, T.; Dey, A.; Ens, B.; Lee, G.; Billinghamurst, M. CoVAR: Mixed-Platform Remote Collaborative Augmented and Virtual Realities System with Shared Collaboration Cues. In *Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017*; Institute of Electrical and Electronics Engineers Inc.: New York, NY, USA, 2017; pp. 218–219. [\[CrossRef\]](#)
21. Bender, S. Headset attentional synchrony: Tracking the gaze of viewers watching narrative virtual reality. *Media Pract. Educ.* **2019**, *20*, 277–296. [\[CrossRef\]](#)
22. Bailenson, J.N.; Blascovich, J.; Beall, A.C.; Loomis, J.M. Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence Teleoper. Virtual Environ.* **2001**, *10*, 583–598. [\[CrossRef\]](#)
23. Keskinen, T.; Makela, V.; Kallioniemi, P.; Hakulinen, J.; Karhu, J.; Ronkainen, K.; Makela, J.; Turunen, M. The Effect of Camera Height, Actor Behavior, and Viewer Position on the User Experience of 360° Videos. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; IEEE: Osaka, Japan, 2019; pp. 423–430. [\[CrossRef\]](#)

24. Speicher, M.; Rosenberg, C.; Degraen, D.; Daiber, F.; Krüger, A. Exploring Visual Guidance in 360-degree Videos. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*; ACM: Salford, UK, 2019; pp. 1–12. [[CrossRef](#)]
25. Pakkanen, T.; Hakulinen, J.; Jokela, T.; Rakkolainen, I.; Kangas, J.; Piippo, P.; Raisamo, R.; Salmimaa, M. Interaction with WebVR 360° video player: Comparing three interaction paradigms. In *Proceedings of the 2017 IEEE Virtual Reality (VR)*, Los Angeles, CA, USA, 18–22 March 2017; pp. 279–280. [[CrossRef](#)]
26. Keijzer, P.D. Effectively Browsing 360-Degree Videos. Master's Thesis, Universiteit Utrecht, Utrecht, The Netherlands, 2019.
27. Petry, B.; Huber, J. Towards effective interaction with omnidirectional videos using immersive virtual reality headsets. In *Proceedings of the 6th Augmented Human International Conference, Singapore, 9–11 March 2015*; ACM: Singapore, 2015; pp. 217–218. [[CrossRef](#)]
28. Elmezeny, A.; Edenhofer, N.; Wimmer, J. Immersive Storytelling in 360-Degree Videos: An Analysis of Interplay Between Narrative and Technical Immersion. *J. Virtual Worlds Res.* **2018**, *11*. [[CrossRef](#)]
29. Rothe, S.; Hussmann, H. Spaceline: A Concept for Interaction in Cinematic Virtual Reality. In *Proceedings of the International Conference on Interactive Digital Storytelling*, Little Cottonwood Canyon, UT, USA, 19–23 November 2019; pp. 115–119. [[CrossRef](#)]
30. Reyes, M.C. Screenwriting Framework For An Interactive Virtual Reality Film. In *Proceedings of the 3rd Immersive Research Network Conference iLRN, iLRN 2017*, Coimbra, Portugal, 26–29 June 2017.
31. Montfort, N. *Twisty Little Passages: An Approach to Interactive Fiction*; OCLC: 799016115; MIT Press: Cambridge, MA, USA, 2005.
32. Marie-Laure Ryan. From Narrative Games to Playable Stories: Toward a Poetics of Interactive Narrative. *Storyworlds A J. Narrat. Stud.* **2009**, *1*, 43–59. [[CrossRef](#)]
33. Kallioniemi, P.; Keskinen, T.; Mäkelä, V.; Karhu, J.; Ronkainen, K.; Nevalainen, A.; Hakulinen, J.; Turunen, M. Hotspot Interaction in Omnidirectional Videos Using Head-Mounted Displays. In *Proceedings of the 22nd International Academic Mindtrek Conference*, Tampere, Finland, 22–24 September 2018; ACM: Tampere, Finland, 2018; pp. 126–134. [[CrossRef](#)]
34. Ibanez, J.; Aylett, R.; Ruiz-Rodarte, R. Storytelling ual environments from a virtual guide perspective. *Virtual Real.* **2003**. [[CrossRef](#)]
35. Sharaha, I.; AL Dweik, A. Digital Interactive Storytelling Approaches : A Systematic Review. In *Computer Science & Information Technology (CS & IT)*; Academy & Industry Research Collaboration Center (AIRCC): Chennai, Tamil Nadu, India, 2016; pp. 21–30. [[CrossRef](#)]
36. Cavazza, M.; Charles, F. Dialogue generation in character-based interactive storytelling. In *Proceedings of the First AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Marina del Rey, CA, USA, 1–3 June 2005; AAAI Press: Marina del Rey, CA, USA, 2005; AIIDE'05; pp. 21–26.
37. Mateas, M.; Stern, A. Towards Integrating Plot and Character for Interactive Drama. *Soc. Intell. Agents* **2006**, 221–228. [[CrossRef](#)]
38. Seif El-Nasr, M.; Milam, D.; Maygoli, T. Experiencing interactive narrative: A qualitative analysis of Façade. *Entertain. Comput.* **2013**, *4*, 39–52. [[CrossRef](#)]
39. de Lima, E.S.; Feijó, B.; Barbosa, S.D.; Furtado, A.L.; Ciarlini, A.E.; Pozzer, C.T. Draw your own story: Paper and pencil interactive storytelling. *Entertain. Comput.* **2014**, *5*, 33–41. [[CrossRef](#)]
40. Cavazza, M.; Lugin, J.L.; Pizzi, D.; Charles, F. Madame bovary on the holodeck: Immersive interactive storytelling. In *Proceedings of the 15th international conference on Multimedia—MULTIMEDIA '07*, Bordeaux, France, 25–27 June 2007; ACM Press: Augsburg, Germany, 2007; p. 651. [[CrossRef](#)]
41. Koenitz, H. Towards a Specific Theory of Interactive Digital Narrative. In *Interactive Digital Narrative*; Routledge: London, UK, 2015; pp. 91–105. [[CrossRef](#)]
42. Signe Brewster. Designing for Real Feelings in a Virtual Reality. Available online: <https://medium.com/s/designing-for-virtual-reality/designing-for-real-feelings-in-a-virtual-reality-41f2a2c7046> (accessed on 14 January 2021).
43. Regenbrecht, H.; Schubert, T. Real and illusory interactions enhance presence ual environments. *Presence Teleoper. Virtual Environ.* **2002**, *11*, 425–434. [[CrossRef](#)]
44. Park, J.N.; Mills, S.; Whaanga, H.; Mato, P.; Lindeman, R.W.; Regenbrecht, H. Towards a Māori Telepresence System. In *Proceedings of the 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, Dunedin, New Zealand, 2–4 December 2019; pp. 1–6; [[CrossRef](#)]
45. Vosmeer, M.; Schouten, B. Interactive Cinema: Engagement and Interaction. In *Interactive Storytelling*; Mitchell, A., Fernández-Vara, C., Thue, D., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2014; pp. 140–147. [[CrossRef](#)]
46. Wei, H.; Bizzocchi, J.; Calvert, T. Time and space in digital game storytelling. *Int. J. Comput. Games Technol.* **2010**, *2010*, 897217. [[CrossRef](#)]
47. Nilsson, N.C.; Nordahl, R.; Serafin, S. Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Hum. Technol.* **2016**, *12*, 108–134. [[CrossRef](#)]