MADE YOU LOOK: AN ANALYSIS OF PERSPECTIVE
AND EMERGING CONVENTIONS OF STORYTELLING IN CINEMATIC
VIRTUAL REALITY

A THESIS
Presented to the University Honors Program
California State University, Long Beach

In Partial Fulfillment
of the Requirements for the
University Honors Program Certificate

Natalie Balderas

Spring 2019
I, THE UNDERSIGNED MEMBER OF THE COMMITTEE,
HAVE APPROVED THIS THESIS

MADE YOU LOOK: AN ANALYSIS OF PERSPECTIVE
AND EMERGING CONVENTIONS OF STORYTELLING IN CINEMATIC
VIRTUAL REALITY

BY

Natalie Balderas

David O’Grady
Film and Electronic Arts
The purpose of this research is to understand what impact the techniques used in cinematic virtual reality have in creating perspective, guiding attention and garnering a sense of immersion. Virtual reality seeks to provide user enhancement through optimized visual and auditory experiences, a goal that parallels the aims of cinema as an immersive storytelling form. There are four main elements of narrative cinema (cinematography, sound, mise en scene, and editing) that are responsible for perceptually cueing and capturing the audience in a captivating story. In a cinematic VR space, these traditional conventions are challenged by user agency. If cinema uses these conventions to manifest plot visually, heighten the sense of immediacy, and create coherent narratives, how do these features manifest in VR narratives? The following analysis of narrative, cinema, and cinematic VR will examine how cinematic VR is attempting to reconcile user agency with traditional storytelling and cinematic visualization.
ACKNOWLEDGEMENTS

I would like to thank my advisor and mentor, David O’Grady, for all the guidance and encouragement throughout this year-long endeavor. I would have been lost without the useful and extensive feedback. Your emails are my favorite; the longer they were the more I enjoyed them. We did it! I am also grateful to my family for supporting me and giving me the resources to pursue my passions. They never question my crazy plans and ambitious projects, but rather cheer me on each step of the way. I owe much of this project to the support from Daniel and Maison. They always listened to my often-nonsensical ramblings and brainstorming, even though neither of them had a strong affinity for what I was talking about. Your feedback helped develop my arguments, ground my ideas, and when I wasn’t making sense, you truthfully told me so. The support and friendship each of you provide is so motivating and gave me the confidence to pursue this research. Thank you. Finally, to the Russian proverb posted on my computer, “There is no shame in not knowing, the shame lies in not finding out.”
TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv

LIST OF FIGURES vi

CHAPTERS

1. INTRODUCTION 7
   Statement of Research Question 7
   Justification for the Research 8
   Definition of Terms 9

2. LITERATURE REVIEW 11
   Narrative, Ergodic Literature, and Game 11
   Immersion and Interactivity 16
   Theories of Cinema and Editing 20
   Virtual Reality and Cinematic Virtual Reality 25

3. METHODOLOGY 28

4. OBSERVATIONS AND ANALYSIS 30
   Narrative Perspective in Literature 30
   Cinematic Perspective 33
   Perspective in Cinematic Virtual Reality 39
   Technique 1: Visual Salience and Selective Lighting 44
   Technique 2: Spatial 3D Sound 51
   Technique 3: Location and Camera Control 54
   Technique 4: Editing 60
   Analysis 63

5. CONCLUSION AND OUTLOOK 70
   Reflection on the Process 70
   Areas for Further Inquiry 71

6. REFERENCES 74
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Trade-offs of each narrative object</td>
<td>19</td>
</tr>
<tr>
<td>2. Film form, the formal system, and stylistic system</td>
<td>21</td>
</tr>
<tr>
<td>3. Walter Murch’s reasons for cutting</td>
<td>36</td>
</tr>
<tr>
<td>4. Visual Salience in <em>Awake</em></td>
<td>46</td>
</tr>
<tr>
<td>5. Selective Lighting in <em>The Great C</em></td>
<td>48</td>
</tr>
<tr>
<td>6. Fixed perspective and lighting in <em>Pearl</em></td>
<td>49</td>
</tr>
<tr>
<td>7. Perspective in <em>Tree</em></td>
<td>50</td>
</tr>
<tr>
<td>8. Camera movement in <em>The Great C</em></td>
<td>56</td>
</tr>
<tr>
<td>9. Camera placement in <em>Tree</em></td>
<td>58</td>
</tr>
<tr>
<td>10. Location control in <em>Awake</em></td>
<td>59</td>
</tr>
<tr>
<td>11. Two taxonomies for guiding attention</td>
<td>65</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

This paper revolves around virtual reality’s status as an emergent media form which can construct narratives. Virtual reality (VR) has garnered much attention recently as many companies have begun developing head mounted displays, controllers, and a plethora of games and experiences for the public. What was previously an academic endeavor only viable on supercomputers in the 1970s and 80s is now accessible on your personal computer. Accompanying this shift is incredible potential for the types of stories told. Storytelling has played such a significant role in progressing society and is the fundamental way in which we make sense of the world. Each story conveys its message or meaning through the sequencing of events, establishing of causal relationships, and the use of time and space in creative ways. With each new medium comes new or modified techniques to best construct an impactful and relevant story. The techniques that best tell textual stories differ from the techniques utilized to best tell highly visual or cinematic stories, though they share similar goals. With the rise of storytelling in virtual reality environments, several opportunities and complications arise in the process of establishing the most effective techniques to accomplish the goals of storytelling. One component that is being negotiated within virtual reality is perspective. As defined for this project, perspective is how the reader/viewer/user is addressed by or inserted into the written, shot, or interactive story. In novels and cinema, conventions were established over time to construct perspective in ways that shape how we perceive the information presented.
and our emotional involvement in the story as it unfolds. However, virtual reality appears to disrupt the stable position of viewer perspective, as VR users now have the ability to look around 360 degrees within the virtual environment and to decide for themselves how they choose to perceive the unfolding story. This emerging development brings me to my research question: What criteria or techniques are constructing perspective within a cinematic virtual reality to guide attention and construct a narrative system, wherein users are explicitly given the freedom or agency to direct their attention and look anywhere within the world? In order to answer this question, or at least begin the conversation about the aesthetics of perspective within a virtual reality environment, an understanding of what traditionally guides and directs attention in literature and cinema is critical.

**Justification for the Research**

Regardless of the medium, consumers want to experience stories that are meaningful and deliver value. Whether that value be life lessons that would otherwise be too painful to learn, the thrill of spectacle and the desire to be entertained, or to facilitate an enjoyable afternoon out with family, we consume stories that offer us something. The establishment of aesthetic criteria and conventions by which narrative is constructed is important in ensuring value is delivered. This analysis seeks to conceptualize the problems facing cinematic VR by utilizing the problem solving of older narrative forms. Viewers do not want to become sick from too much motion, disoriented from jagged cutting, or overwhelmed from a too intensive soundscape: any of those factors would discourage consumers from dropping the significant amount of money necessary to run a VR system. When viewers see their desires are fulfilled, they are more likely to make the
investment. In order to develop the medium, both experimentation within cinematic VR and determining the criteria that cinematic virtual content can be held accountable to are necessary. Often, experimentation leads to the establishment of criteria, as famously epitomized in cinema by Edwin S. Porter’s *The Great Train Robbery* (1903). The artists and studios were not surveying viewers about how to inject narrative into the cinematic spectacle, but rather they had to discover through trial and error that narrative could be progressed through the splicing of film strips together. Thus, editing emerged as a convention for telling stories. Today, with each new cinematic VR short created, valuable lessons can be learned which collectively paves the road for mainstream adoption and integration of virtual reality as a storytelling vehicle. With each new cinematic virtual reality short content creators do not have to reinvent the wheel. By evaluating the role of user perspective in current popular cinematic VR shorts, we can begin to establish some formal techniques that cinematic VR can employ to reconcile user freedom or agency with visual storytelling.

**Definition of Terms**

*Attention:* Attention is the perceptual process of singling out and focusing on specific information from the environment.

*Cinematic Virtual Reality (cinematic VR):* A term deemed by Oculus Creator Studio to represent the genre of virtual reality content that constructs narrative stories for viewing in virtual reality technology.
**Perspective:** Perspective is the frame or lens through which the story, world, and characters are observed. The creator utilizes perspective in order to best present information to the reader/viewer/audience.

**Immersion:** Immersion can be understood as a perceptual shift in which viewers willingly choose to insert themselves in a created world and to accept the happenings within that world as a persuasive simulation of reality.

**Interactivity:** The ability to impact the environment and engage elements within the constructed media object which have a noticeable result on the progression of story, or which modify the construction of meaning.
CHAPTER 2

LITERATURE REVIEW

Narratives, Ergodic Literature, and Games

One of the most impressive and effective methods of communication throughout human history has been the story. Fundamental messages, themes, and advice are conveyed through the successful adoption and sharing of stories from generations of communities and cultures. Stories touch deeper inclinations we have about the world, they provide guidance in our individual journeys through life, as seen through Joseph Campbell’s critical studies on narratology: “Throughout the inhabited world, in all times and under every circumstance, myths of man have flourished; and they have been the living inspiration of whatever else may have appeared out of the activities of the human body and mind. It would not be too much to say that myth is the secret opening through which the inexhaustible energies of the cosmos pour into the human cultural manifestation” (Campbell 1). Emphasized in this quote is the idea that myth, an early and enduring form of storytelling, is not simply a form of entertainment, but a method by which we structure and perceive our world.

Myths not only shaped our perception of the world, they also established a memorable and relatable form—the story—to impart information and wisdom. Transcending the content and form of each individual story is the overarching structure throughout all stories, understood as narrative. In his book *Narrative Comprehension and Film*, Edward Branigan defines the ways in which narrative functions: “Narrative is a
perceptual activity that organizes data into a special pattern which represents and explains experience. More specifically, narrative is a way of organizing spatial and temporal data into a cause-effect chain of events with a beginning, middle, and end that embodies a judgement about the nature of the events as well as demonstrates how it is possible to know, and hence to narrative, the events” (Branigan 3). Narrative is a broad category that encompasses many forms and types of stories, both fictional and real. For example, we are constantly constructing stories and organizing the events and information of our daily lives into cohesive, even dramatic, stories that we share with family and friends.

With the advent of personal computers and the digitization of nearly all cultural forms, including storytelling, new story forms were created that challenged the conventions of narrative, including the role of the observer, the perspectives presented, and the overarching structure of narrative itself. Specifically, computers opened the door for the development of interactive narratives, where users play a role in progressing the story forward and shaping the outcome. These types of stories integrate the element of choice into them, and much of the hassle of facilitating choices or managing forking paths—as seen historically in both experimental and popular literature, such as Choose Your Own Adventure novels—are smoothed, polished or hidden by computer code. Literary theorist Espen Aarseth refers to such interactive narratives as cybertexts or ergodic literature. Both terms are based on the reader in the narrative experience, and suggest an elevated level of performance, or work, that is required for these readers to progress the story. Cybertext and the broader classification of ergodic literature has shaped the modern understanding of and goals for digital literature:

The concept of cybertext focuses on the mechanical organization of the text, by positioning the intricacies of the medium as an integral part of the
literary exchange. However, it also centers attention on the consumer, or user, of the text, as a more integrated figure than even reader-response theorists would claim. The performance of their reader takes place all in his head, while the user of cyber text also performs in an extranoemetic sense. During the cybertextual process, the user will have effectuated a semiotic sequence, and this selective movement is a work of physical construction that the various concepts of “reading” do not account for. This phenomenon I call ergodic. In ergodic literature, nontrivial effort is required to allow the reader to transverse the text. (1)

This term suggests that rather than internalizing the comprehension process, externalized effort is required to get the full depth and breadth of the narrative content. These theoretical classifications hint toward a growing trend in the digital age of personalization and user control, which are reflected in the types of stories crafted with new media, where users must work to progress the text. The result is a self-catered story wherein the “reader” plays a significant role in story creation and whose participation is an essential factor in the form. Interactive storytelling is so culturally pervasive today—as seen in videogames and interactive stories developed in Twine, for example—that even conventional media is experimenting with it. For example, the recent Netflix hit Bandersnatch, released in December 2018, follows a young programmer developing a computer game. Throughout the film, the viewer is prompted to choose between two actions to progress the story; some responses continue to advance the story, but others may result in the protagonist’s death.

Critical to understanding the development of virtual reality as an interactive medium is an appreciation for narrative’s counterpart, the game. Both literary and games theory have explored the complicated intertwining of the two media as a result of technological advancement; noting the distinctions and critical components which comprise interactive narrative and games is essential for understanding virtual reality’s
dependence on both media. As Aarseth states, “Cybertext is not a new, revolutionary form of text, with capabilities only made possible through the invention of the digital computer. Neither is it a radical break with old-fashioned textuality, although it would be easy to make it appear so,” (18). Much like the cybertexts that came before it, VR is not a radical departure but dependent on resolving many of the same issues of interactive storytelling.

Media theorist Britta Nietzel provides an appropriate definition of the logic of games, where “…a state of being is maintained, the result is not yet decided, and one is still in the process of fixing an outcome. In this sense, we can describe games as a process of expanding the present, in which certain defining tendencies for the future are laid down by the rules, not the exact chain of events or the results. Games cannot be adequately described without considering the process of play” (Nietzel 229). When a game is played, the route to the ending varies depending on the player’s actions and decisions which influence interpretation and progress. This suggests one primary difference between game and narrative, and that is the role of agency granted to the player/reader. Jesper Juul provides a greater distillation of this difference in his essay, “A Clash Between Game and Narrative”:

It does seem reasonable to claim that narratives are sequences evoking a sense of destiny, of events that have to lead to each other. Roland Barthes says that narrative is the language of destiny, "the mainspring of narrative is...what comes after being read in narrative as what is caused by". Sequence matters in narratives, and the famed translatability of narratives between different media does presuppose fixed sequences. Unlike the fixed sequence of the narrative, games seem to be based upon the relative freedom of the player, on the players’ possibility of influencing the course of action. (Web)
Whereas narratives offer limited agency through the construction of a fixed sequence, games provide relative freedom by using a flexible sequence. Interestingly, this once clear distinction has blurred with the development of computers. We can refer to Aarseth’s conception of ergodic literature in which users have to put in nontrivial work to progress the narrative, a notion that hints at a greater integration of agency and relative freedom into various narrative structures. There are many elements of game that supplement and add to the success of narratives, and vice versa.

Rather than relying on external factors presenting information in an orderly sequence, the primary form of games is to present the world to the player through the perspective of a character or camera which you have control over. Nietzel provides insight to the idea of perspective within videogames: “Apart from the situations in which characters of the computer game speak, players gain all knowledge in computer games through the things they can see, the perspective in a spatial sense. This perspective in the computer game influences the distance of the player to the virtual world,” (238). As a result, the player can look around and explore their environment as though part of the world, and in this sense the perspective is personal and individual. No two players will consume the content the same, different players will direct their attention towards different elements of the game, yielding different, albeit similar, perceptions of the story. This is challenged with the adoption of cutscenes and fadeaways in newer games in which player control is taken away to present information relevant to the story, however this is only further testimony about the merging of media forms. The fundamental factor stands; through videogames we can observe a thorough integration of first-person perspective into stories, wherein a subjective perspective is utilized to create a more
immersive experience. This is the tantamount goal of virtual reality: presenting a story and awarding agency to a viewer through a uniquely personal and perceptual experience. Attempts to embrace player agency while simultaneously presenting narrative structure is resulting in new forms and techniques, along with potential paradoxes in striving for user interactivity while maintaining a coherent structure.

On Immersion and Interactivity

As the development of virtual reality and augmented reality have garnered popular appeal and media attention, so has the idea of interactivity. Theorist Marie-Laure Ryan defines interactivity as:

… not merely the ability to navigate the virtual world, it is the power of the user to modify this environment. Moving the sensors and enjoying freedom of movement do not in themselves ensure an interactive relation between a user and an environment: the user could derive her entire satisfaction from the exploration of the surrounding domain. She would be actively involved in the virtual world, but her actions would bear no lasting consequences. In a truly interactive system, the virtual world must respond to the user’s actions. (121)

Here, we begin to see the relationship between user agency and interactivity in that nontrivial effort is required to transverse the text and elicit responses from the environment, a fundamental distinction noted previously by Aarseth in classifying ergodic literature. Lev Manovich additionally provides context and criticism for the term, specifically for its lack of defined characteristics: “Modern HCI (Human-computer Interfacing) is by definition interactive. In contrast to earlier interfaces such as batch processing, modern HCI allows the user to control the computer in real-time by manipulating information displayed on the screen. Once an object is represented in a computer, it automatically becomes interactive. Therefore, to call computer media
‘interactive’ is meaningless— it simply means stating the most basic fact about computers” (55). Despite this criticism on the use of interactivity for the vague and nonspecific nature of the idea, he does go on to suggest a type of interactivity that is more complicated, and that emphasizes the ways users interact with these objects. Rather than equating the idea of interactivity with a literal physical interaction, we may observe interactivity as a more “psychological process of filling-in, hypothesis formation, recall, and identification, which are required for us to comprehend any text or image at all,” (Manovich 57). Interactivity can be observed to varying degrees and with differing sets of criteria, however the consistent motif of interactivity is that it is used to engage and activate the user.

The true risk of allowing the user unmitigated control, or interactivity, within the environment is the increasing impossibility to maintain a semblance of narrative coherence. In fact, returning to the idea of cutscenes in videogames, Ryan suggests that, “…the fact that it is necessary to temporarily remove control from the user to establish the narrative frame is a further indication that interactivity is not a feature that facilitates the construction of narrative meaning” (Ryan 602). This is where the paradoxical relationship between narrative and interactivity arises, though, like the distinction between game and narrative, the landscape is currently being negotiated as new forms are emerging.

Another critical concept, immersion, must be acknowledged in order to understand the extent to which this negotiation is occurring. Immersion is best equated to the term, suspension of disbelief, to quote Pimentel and Texeira in Ryan’s article:

As [users] enter the virtual world, their dept of engagement gradually meanders away from here until they cross the threshold of
involvement. Now they are absorbed in the virtual world, similar to being in an engrossing book. The question isn’t whether the created world is as real as the physical world, but whether the created world is real enough for you to suspend your disbelief for a period of time. This is the same mental shift that happens when you get wrapped up in a good novel or become absorbed in playing a computer game. (Ryan 114 from Pimmentel and Texeira 15)

Immersion can be understood as a perceptual shift wherein the viewer or reader willingly inserts them self in the created world and accepts the happenings within that world as a reality. This suggests that immersion can occur without interactivity because the viewer is not required to engage with the work beyond looking, listening, and imagining the world that is being presented to feel immersed.

So, whereas interactivity is the extent to which a user can impact the progression of the story as well as alter their perspective, immersion is the willingness of the user to suspend their disbelief so that they can experience the full emotional and contextual experience of the narrative. Certain media forms facilitate the process of immersion at a quicker pace and make the process easy, whereas others require more of an investment from the viewer in order to achieve immersion. A classic novel from a writer with a distinct dialect requires time to adjust before the reader can feel fully immersed in the world and happenings, whereas a film depicting the same classical story can be understood quick due to the highly visual capabilities of cinema. Again, as seen with interactivity, there are trade-offs between immersion, user agency, and impact of the narrative. In order to achieve the greatest degree of immersion, certain conditions for each medium should be met, which requires work on the part of the reader/user. For example, in order to properly immerse yourself in a movie, it is recommended you watch in a dark, quiet environment with a large screen. This does not mean that you must watch
a movie this way, however immersing yourself in a film is much more difficult on your iPhone in a crowded train station.

Additionally, immersion and interactivity do not always lend themselves to one-another. Marie-Laure Ryan provides a helpful model of comparison which juxtaposes immersion, interactivity, and global design to illustrate the inevitable trade-offs that accompany each type of media object. In figure 1, this negotiation is exemplified and highlights the related but different emphases of each object. She goes on to explain her scores:

The strong forms of interactivity run most blatantly into the problem of design: how can the contributions of the reader-turned-author be monitored by the system, so that the text as a whole will maintain narrative coherence and aesthetic appeal? In the weaker forms of interactivity, design is easier to control, but immersion remains problematic. The various attempts by contemporary literature to emulate the interactivity of VR create a loss of involvement in the fictional world, a weakening of the imaginative experience and a momentary breakup of make-believe. (Ryan 132)

While these can be restrictions to the types of stories told, the media in which they’re told, and the ways in which perspective is constructed, they can also be viewed as opportunities to craft and redefine new cinematic narrative experiences in virtual reality.

---

Immersivity: 1. Traditional narrative. 2. Role-playing games. 3. Hypertext.

Interactivity: 1. Role-playing games. 2. Hypertext. 3. Traditional narrative.


---

Theories of Cinema and Editing

Cinema is a spatial-temporal, photographic medium that comes into contact with narrative cause and effect. The medium represented a merge of previous storytelling vehicles such as theater, visual novels, and photography into a grand spectacle on the big screen. Film is a particularly immersive form, something Andre Bazin refers to as presence and which is explained further in Steven Katz’s book on film directing:

Presence is used to describe the moviegoer’s sense that he is within the same spatial/temporal continuum as the picture on the screen. He (Bazin) regarded this illusion as the fulfillment of a tradition of verisimilitude in western painting that began with the discovery of linear perspective during the Renaissance. Motion pictures take this one step further...instead of seeing the picture surface we are included in the pictorial space projected on the screen as if it were real three-dimensional space. (3)

Movies present narratives in a deeply immersive way; the large screen dominates our vision, the dark room directs our attention, and the encompassing audio places us in the world of the story. The ability to forge great connections with audiences through visual and auditory stimulation while minimizing exposure to any inherent risk lends to cinema’s ability to tell immersive, impactful, and entertaining narratives.

Our ability to comprehend the sequential presentation of images as motion, and to extract meaning from what we are watching gives testimony to the natural tendency of our brains to construct narratives. Lev Manovich goes so far as to suggest that, “cinematic means of perception, of connecting space and time, of representing human memory, thinking, and emotion have become a way of work and a way of life for millions in the computer age” (Manovich 86). What used to be conceived as mere “cinema of attractions,” wherein film played the role of commodity and spectacle, has stemmed into a multi-billion dollar industry that crafts stories for mass audiences.
Further, cinema redefined the critical relationships between viewer and author through the use of the camera and did so not by adhering to a fixed point of perspective, as found in novels, or a fixed viewing distance, as found in theater.

Even before the mainstream adoption of cinema as a significant artistic, cultural, and storytelling form theorists began assembling aesthetic criteria by which film content can be evaluated. David Bordwell and Kirstin Thompson are two notable authors and theorists who have written extensively on film form, narrative as a formal film system, and film’s stylistic system as the “patterned and significant use of techniques,” (Bordwell and Thompson 327). Film’s form is a composition of its formal and stylistic systems interacting with one another to create and depict the progression of the story (see fig. 2). They state, when elaborating on film style:

The spectator has a relation to style as well. Although we are seldom conscious of the fact, we tend to have expectations about style. Like other kinds of expectations, stylistic ones derive both from our experience of the world generally and our experience of film and other media. The specific film’s style can confirm our expectations, or modify them, or cheat, or challenge them… Thus the filmmaker’s technical decisions make a difference in what we perceive and how we respond. (Bordwell and Thompson 328-29)

Style plays a significant role in guiding our attention and developing the perspective, and this formula suggests that with each new form, the stylistic techniques adapt to meet the demands of whatever formal system is being articulated. Bordwell and Thompson declare four primary stylistic techniques which play a significant role in the construction of the filmic form. As previously noted, they are mise-en-scene, cinematography, editing, and sound, and their contributions are what distinguish cinema from other media forms, and how each film differs from each other. As part of my methodology I will be analyzing how each cinematic VR short conforms to these pivotal cinematic stylistic techniques to understand the ways cinematic VR as a new media form is adapting to and changing previously grounded cinematic conventions, and why. For the purposes of this thesis I am focusing my discussion on how these conventions serve in relation to editing, though the analysis section will have much crossover.¹ Editing is the site of the most tension within cinematic VR, where editing previously played the most significant role in constructing perspective, guiding attention, and garnering immersion.

Before proceeding, establishing the definitions and utility of mise-en-scene, cinematography, and sound will provide greater insight when conducting analysis later. Mise-en-scene is a broad term which encompasses all the visual information provided in the screen. To quote Bordwell and Thompson: “As you would expect from the term’s theatrical origins, mise-en-scene includes those aspects of film that overlap with the art of theater: setting, lighting, costume, and the behavior of the figures(actors),” (156). This is

¹ For those interested in delving into mise-en-scene, cinematography, and sound in further depth, I highly recommend reading Bordwell and Thompson’s book Film Art: An Introduction.
the primary means through which the director’s style is conveyed and is often the most recognizable feature of a film. Second, cinematography involves the stylistic creation and manipulation of the camera to create the shot. Again, returning to Bordwell and Thompson, “the filmmaker uses the camera to regulate how light from some object will be registered. In any event, the filmmaker can select the range of tonalities, manipulate the speed of motion, and transform perspective,” (193). Because the cinematographer chooses the frame and area to look at within the film, they exert immense control over how the story is perceived, and how information is revealed. Third, sound in cinema plays a critical role in orienting the viewer, adding complexities, and building the world. The use of sound to alter perception is another significant contribution, with tools such as loudness, pitch and timbre contributing to the sense of being immersed in the film. Each of these stylistic techniques work together and accomplish the same goal: to immerse the viewer in the story and the message. Their usage varies depending on the genre and artistic input, however together each aspect contributes to our understanding and acceptance of film’s form.

In cinema, the motion picture editor must take all the individual shots and compile them to forge a coherent story—a responsibility which weaves all elements together into a beautiful, emotional, and unified whole. Lev Manovich states: “Editing, or montage, is the key twentieth-century technology for creating fake realities. Through editing, images that could have been shot in different geographical locations or at different times create the illusion of a contiguous space and time,” (148). Through much technological development and experimentation, cinema editors have adopted a critical role, highlighted by Tom Rolf in his statement, “All movies are really four movies: the
movie imagined, the movie written, the movie directed, and the movie edited” (Tucker 91). This elevates the role of editing to a position equal to that of the author or director, capable of impact not previously recognized or granted to editors in other narrative forms. Editors of written texts, while playing an important role in getting the work ready for publishing, are by no means authors of the narrative time and space of the work. That role of establishing and creating the story is delegated to the author, the editor is merely a second pair of eyes offering advice.

The impact editing has on the construction of successful cinematic narratives lead to the development of certain aesthetic and technical criteria for effectively weaving separate shots into a unified whole. Best practices for editing were established early on for maintaining narrative coherence and solving the spatial-temporal problems arising from storytelling such as parallel narratives is known as classical editing, or continuity editing. This style became the predominant means for constructing film narratives as early as the 1920s and is still relied upon today. Bordwell and Thompson provide context to the style:

Filmmakers came to assume that a film should guide the spectator’s attention, making every aspect of the story on the screen as clear as possible. In particular, films increasingly set up a chain of narrative causes and effects. One event would plainly lead to an effect, which would in turn cause another effect, and so on...Every aspect of the film style came to be used to enhance narrative clarity...Editing came to emphasize continuity among shots. Certain cues indicated that time was flowing uninterruptedly across cuts. Between scenes, other cues might suggest how much time had been skipped over. When a cut moved from one space to another, the director found ways to orient the viewer. (Thompson and Bordwell 43)

Film editing today generally adheres to this formula, however there is certainly experimentation as filmmakers attempt to achieve faster cutting, more dramatic emphasis, or surreal representations of reality.
Virtual Reality and Cinematic Virtual Reality

Researchers and theorists have grappled over the precise definition of virtual reality for decades, beginning with early development in the 1960s and 70s to the newest devices of the last decade. Howard Rheingold, a well-known American critic of digital technologies, specified two foundations of virtual reality in 1991:

The idea of immersion - using stereoscopy, gaze-tracking, and other technologies to create the illusion of being inside a computer-generated scene - is one of the two foundations of VR. The idea of navigation-- creating a computer model of a molecule or a city and enabling the user to move around, as if inside it -- is the other fundamental element. (112)

Working in tandem with Rheingold’s definition are the technical aspects which render immersion and navigation within VR possible. From a technical standpoint, The Jaunt Team, a VR production company, have four specifications for virtual reality, “360° Equirectangular Images, Stereoscopic 3D, Spatial 3D Sound, and Viewed in a Head-Mounted Display,” (Jaunt VR Studios). A third, and more vague approximation of virtual reality is provided by Jason Jerald in his book titled The VR Book: “Virtual reality is defined to be a computer-generated digital environment that can be experienced and interacted with as if that environment were real” (Jerald 9). Each takes a different approach toward their definitions, however each provide relevant information regarding virtual reality’s premise. They describe the creation of a space that Myron Krueger describes in his book Artificial Reality, known as a responsive environment, which is “an entity that engages participants in dialogue, a multidimensional space which the participant can explore, and where each subspace has unique response relationships, an experiential parable where the theme is illustrated by the things that happen to the protagonist - the participant,” (Krueger 46). This environment fosters audiences who are
active participants in the space around them, rather than passive viewers\(^2\). Virtual reality has applications in many fields, from medicine to aerospace to the arts. In particular, virtual reality as an entertainment medium has risen in popularity with the introduction of Oculus Rift and Google’s Cardboard in 2016, and subsequent head-mounted displays released by Lenovo, HTC, and Sony. In these environments, viewers wear the headset that is motion tracked by cameras for precise head and arm movements. 3D and 360-degree immersion is achieved by streaming independent images to each eye at a high refresh rate, thus creating a strong visual illusion of presence. With the audio, viewers wear headphones that project stereoscopic sound to deepen the 3D space and add depth cues.

Emphasis in marketing is directed towards VR’s ability to provide immersion, navigation, and interaction with a responsive environment. When integrated properly, virtual reality can elevate what videogames and cinema each do individually to a new level. Jaunt Studios, a VR production company, defines two major realms within which VR as entertainment occupies: games--coined game-based virtual reality--and narratives-defined as cinematic virtual reality. John Mateer, in his article on directing within cinematic VR, defines the term as, “a type of immersive VR experience where individual users can look around synthetic worlds in 360°, often with stereoscopic views, and hear spatialized audio specifically designed to reinforce the veracity of the virtual environment,” (Mateer 15). However, whereas VR allows for direct manipulation and alteration of objects in the environment, cinematic VR is rather limited in scope: “the

\(^2\) The use of the term passive is not to suggest that readers, viewers and users of non-interactive media are not experientially active, but rather to indicate that the environment, or mise-en-scène, is not responsive to viewer input or manipulation in traditional storytelling media.
inability of users to actually interact with elements contained within the cinematic virtual world is the primary difference between the two media. While both are immersive, cinematic VR experiences are effectively linear presentations with the duration of each experience dictated by the length of the media assets employed,” (Mateer 16). Jaunt VR defines cinematic VR as, “360° video filmed using a panoramic video camera system and played back as an equirectangular video file which allows the user to look around the scene as it unfolds,” (Jaunt VR Studios 9). Cinematic VR is an experience wherein a visual story is presented in 360 degrees, giving the viewer the ability to look anywhere within that constructed space. While they additionally note the limitation of cinematic VR in regard to viewer interactivity, they nod to opportunities for growth: “Though not as interactive as full game-based environments you can still add interactivity to cinematic VR. Branching “Choose Your Own Adventure” stories, gaze detection, interactive overlays and interfaces, audio or clip triggers, gestures, and even full CG integration are all possible,” (Jaunt VR Studios 9). While viewers have agency to choose where they look within the field, they cannot significantly alter the story itself. This places limits on both the interactivity and responsiveness of the environment, however, hardly limits the creative construction of unique experiments which are attempting to push those boundaries.

In conclusion, this literature review suggests opportunities to explore the ways cinematic VR is attempting to provide an immersive and interactive narrative experience. Through an understanding of the problems and solutions that have arisen and been reconciled in previous narrative forms, we can gain a new perspective on the issues of perspective within this emerging medium.
CHAPTER 3

METHODOLOGY

The goal of this thesis is to explore the emerging techniques for constructing perspective in cinematic virtual reality. This study is comprised of research into the appropriate academic literature and an analysis of several media objects. Literature in the fields of narratology, cinema, and virtual reality establish a framework for understanding the theories and practical conventions which are responsible for and interact with our understanding of how perspective can be shaped in cinematic VR. The paper conceptualizes how traditional methods of visual storytelling and cinematic narratives are being renegotiated to operate within the space of a growing technology, virtual reality.

With the research portion, academic articles, books and literature from the fields of virtual reality, narratology, games theory, cinema studies, and new media are integrated to establish a frame of understanding. This provides insight into the complications and considerations associated with cinematic virtual reality. Further, the research will reveal the formal characteristics of cinema that encounter virtual reality, in particular, issues of viewer interactivity, cinema-editing, and directing user perception. Additionally, the research will engage theories of narratology and game, with the purpose of exploring the ability to tell stories where interactivity is a critical component of virtual reality entertainment.

Additionally, this paper will conduct several media content analyses of cinematic VR shorts as a way of gauging the effectiveness of current interactive experiences and
perceptual cues. Each media object, comprising of various cinematic virtual reality experiences, were selected on the basis of several criteria, chiefly, for how they construct perspective, how they draw upon past conventions of cinema, the responsibilities each grants the viewer, how they draw attention to the relevant parts of the environment, and how they maintain immersion. The four films which will be referenced in the observation and analysis section are The Great C (2018), Awake (2018), Tree (2017), and Pearl (2016). These cinematic VR short films were selected as a representative sample of high profile, mainstream examples. They were also selected because they exhibit a wide range of perspective techniques and methods for reconciliation. Looking to various visual narratives that have emerged in virtual reality offers tangibility to the theoretical research and suggests ways in which the relationship between audience and creator are being negotiated.

The predominant limitation with the methodology and more broadly the thesis is that the focus is on an emerging media art form that has yet to solidify in culture. Therefore, the thesis and research cannot make any definitive claims about the future of this technology or long-term consequences relating to culture, industry, and entertainment. The goal of this paper is to understand the challenges cinematic VR faces in creating interactive and immersive narratives. In looking towards academic literature in the fields of narratology, cinema, and new media, an enhanced understanding will emerge of cinematic virtual reality’s origins, along with medium-specific issues of viewer perception, interactivity, and editing. The information from the content analysis will provide greater context and understanding of the issues, as well as parse through the solutions and opportunities the virtual industry is creating for virtual reality.
CHAPTER 4

OBSERVATION AND ANALYSIS

The context and research provided in the literature review lays the groundwork for a greater understanding of how each media form constructs the viewer’s perspective and perception of the events transpiring in a narrative system. Further, this leads into a discussion regarding how cinematic VR inherits the tools of perspective in narrative literature and cinema and problematizes those conventions. The narrative system is comprised of a progression through space and time to present a causal chain of events that come together to convey a message and experience. Without a framework for conceptualizing and viewing, you are just observing without motivation for emotional involvement. Perspective is a stylistic technique which properly orients the reader or viewer and constructs the narrative so that they can extract meaning and information from what they are observing. One of the main questions that will be raised and answered in each section of this analysis is, “How is the reader/viewer positioned in relation to the text, and how is that perspective shaped, directed or guided?” Each form has different needs and demands, and as a result handles these issues in different ways.

Narrative Perspective in Literature

The perspective, or focalization, in literature has a significant impact on the information revealed to the reader. Focalization is a process of foregrounding certain elements and characters which are deemed most significant by the author, and thus most relevant for the reader to pay attention to. In her chapter on “Narrativity in Computer
Games,” Britta Neitzel distills the theory of perspective, point of view, or focalization, as developed by Gerard Genette:

These aspects of the mood apply to the extent of knowledge that the discourse will reveal about the narrated world. He defines three types of focalization consisting of: zero focalization or the non-focalized (objective perspective) story, i.e., a story that employs an omniscient narrative; internal focalization (subjective perspective), in which one or more people in turn narrate from their own viewpoints; and external focalization (semi-subjective), in which the viewpoint concentrates on the actions of one person, although nothing about the feelings of this person is revealed. (238)

From this distilled classification of perspective arises a large swath of ways in which stories are focalized, including first-person, first-person plural, second-person, third-person, omniscient, and limited omniscient. Interestingly, and to Genette’s point, in literature authors often employ several different perspectives at different points in the story. Each of these focalization points serve as a tool to shape the reader’s perception of the events that are transpiring. Perspective can be used in many ways, such as to get you to empathize with certain characters, despise others, anticipate an outcome, or leave you in suspense wondering what will happen next.

The concept of focalization as a stylistic technique in literature, though not only visible in literature, provides an answer to the question of how our looking is directed. We are being presented with a frame through which the world is observed that emphasizes certain events, characters, or actions over others, thus directing our focus to what the author designates as important. Through the author’s construction of the world we are directed, the author exerts immense control over what we are looking at. Authors make us look through the construction of each word, sentence, and chapter, which are written with intention and purpose. All the information we need is provided and
conveyed in the text, all the details are meant to cue us into the story and maintain our attention. The author may choose to present the information in a subtle manner to underplay its significance, or build suspension, passing responsibility onto the reader to interpret and perceive those cues.

In this regard, an aesthetic goal of perspective in literature is understood as a deliberate presentation of the world for the reader. Authors use perspective to orient or frame the reader in relation to the narrative, as well as to underscore the narrative’s aesthetic efforts to immerse the reader in a story that may be dramatic, comedic, romantic, etc. This is what Gerard Genette is referring to when he uses the term mood. For example, in a dramatic story an author might employ a more objective focalization so that the reader knows more than the characters in the world, thus achieving elevated suspense and dramatic emphasis. This may contrast with a romantic story, which utilizes several subjective perspectives to immerse the reader in the love experienced by each character for the other.

The role of the reader is therefore the imaginer, responsible for taking the words on paper and constructing the world in their mind. This process is an immersive one, and though the author has technically constructed the world, the reader must recreate that construction so that they can understand the story. The reader is also responsible for suspending their disbelief and accepting the world on paper as a reality worth observing. If the sentences are sloppy, descriptions vague, and characters unrealistic, the reader will have a difficult time suspending their disbelief to immerse in the story. In ergodic literature and games, the reader/player has the additional responsibility of making non-trivial actions to progress the story. Here, the reader is granted some control over the
information presented to them, suggesting a renegotiation between the role of the author and that of the reader. In a way, the reader becomes the author of their own read-through, making decisions which ultimately impact their perception of the story. This is also the case for narratives that experiment with and attempt to break down the cause-and-effect chain of events, such as the Choose Your Own Adventure novel or hypertextual works. By granting the reader control over the decisions made in the world, they become an active participant in shaping the perspective, or focalization, of the story.

**Cinematic Perspective**

Cinema, through its expressed ability to immerse viewers in the spatial and temporal continuum presented on screen, heightens stylized perspective to a new level. Rather than maintaining a fixed viewing distance or focalization point, cinema can manipulate the viewing space to construct a highly immersive world on the screen. The ability to construct perspective through methodical shot composition and editing facilitates a new relationship between the viewer and how they look at the world. Thanks to the formal properties of cinematography and editing, many new stylistic techniques associated with directing attention arose and are recognized as conventions of the filmic form. This section will discuss the ways perspective is constructed in cinema, and how this impacts the role of the viewer and their ability to perceive the cinematic world. Further, we will look at how the four critical stylistic techniques highlighted by Bordwell and Thompson contributes to the construction of perspective in cinema.

Perspective is primarily shaped through the placement of the camera within the cinematic world; viewers are afforded the ability to look into the onscreen world dependent upon how and where the camera is situated. This relationship has led to the
dominant metaphor developing in the 20th century of the viewer as the “camera eye.” Wherever the camera is located is where the viewer is placed in the scene to observe the action. Through the conscious decisions to move the camera in close or observe from afar, filmmakers direct the audience's attention and shape their perception of the events unfolding. There are several types of perspective presented in cinema, and they are all dependent on the use of the camera. While most films present a fairly ‘objective’ omniscient camera, which merely observes and plays no role within the film, a subjective camera can be used to heighten emotion, maintain dramatic tension, or reveal information about the states of characters. The cinematic camera can never be purely objective though, as the filmmakers selectively filter the information observed. However, the title of objective camera in this sense refers to the idea that the camera does not play a role within the story world, and is a non-diegetic element, therefore not subject to the rules and subjectivity of that world. The separation between the story and camera allows the filmmakers great control over how the story is presented and the focalization of each action that occurs. Cinematography is a significant resource for constructing perspective, and various tools such as focal length, depth of field, speed of motion, framing, and camera placement heavily impact the visual creation of perspective.

Mainstream cinema has strayed away from constructing the story in a consistent first person or subjective perspective wherein the viewer embodies a character in the world, such as is the case in many videogames. The general absence of literal character perspective, also known as subjective shots, is a result of the plethora of identification issues that accompany attempts to frame the camera within the head of a character. Media
theorist Alexander Galloway discusses the scarcity of subjective shots in his book *Gaming: Essays on Algorithmic Culture*, stating their narrow range of functions:

Most narrative films don’t include a single subjective shot, and in the films that do, there are generally only a handful of subjective shots used to achieve very specific results. When a subjective shot is used, it generally signifies some type of negative vision. It is sometimes an evil vision, or an inhuman one, or simply a moment of alienation or detachment within a character. (Galloway 46)

If filmmakers want a more regular or objective view of things, they opt for point of view shots rather than subjective shots. Even so, POV shots, where the camera objectively shows us what a character is looking at, function as heightened or stylized moments, and are not the primary means for conveying the story. In videogames, players are often granted the ability to control where their “camera eye” looks and moves, and therefore the sense of immersion amidst the highly subjective and personalized perspective is maintained. Granting the player the ability to control where they look works because they must navigate the game, which ultimately influences the way they perceive the story and world. This requires that the cinematic authors, especially the cinematographer, surrender control over some aspects of the stylistic techniques. Currently, the conventions of the cinematic form favor authorial control; maximum immersion is achieved through the decisions of the filmmakers.

The second critical element which crafts perspective and maintains the focalization in cinema is editing; the sorting and ordering of shots to tell a story in mainstream filmmaking. Each sequence is constructed intently, with the goal of establishing the meaning, story, and emotion for audiences.
Influential film editor Walter Murch has written extensively about the relevancy of film editing to audience engagement. Murch proposed a hierarchical structure for organizing the motivations that informs editing decisions to cut and the overall significance cutting has on all elements of the film form, known as the rule of six. In his hierarchy, presented in figure 3, the reasons are ranked in order of significance. This structure is explained in his seminal work *In the Blink of an Eye*: “Emotion is worth more than all five of the things underneath it. And, in fact, there is a practical side to this, which is that if the emotion is right and the story is advanced in a unique, interesting way, in the right rhythm, the audience will tend to be unaware of (or unconcerned about) editorial problems with lower-order items like eye-trace, stage-line, spatial continuity, etc.” (Murch 19). In other words, editing’s primary function is to present perceptual cues which support the emotional impact of the narrative; therefore the viewer’s looking is directed through cinema’s ability and expertise in emotionally impacting the viewer. This emphasis gives credence to the immense power of the suspension of disbelief and suggests that technical blunders and continuity errors will be overlooked when the viewer is captivated by the emotionality of a good story.

*Fig. 3. The reasons for making a cut. Murch, Walter. In the Blink of an Eye: A Perspective on Film Editing. 1st ed. Los Angeles: Silman-James, 1995. Print.*
Another element of cinema that significantly impacts the way perspective is constructed is the presence and persistence of the screen. To quote Lev Manovich: “The visual culture of the modern period, from painting to cinema, is characterized by an intriguing phenomenon --the existence of another virtual space, another three-dimensional world enclosed by a frame situated inside our normal space,” (95). This idea is a progression of the concept of western perspective and framing developed by Leon Battista Alberti, an Italian Renaissance artist and author writing in the 1400s, where a rectangular canvas acts as a window into another space. With this rose the use of a vanishing point to construct perspective paintings in order to replicate our eyes’ ability to perceive our world, so that the world inside the canvas appeared to be beyond the frame. As a result, a newfound relationship arose between the viewer and screen: “Although the screen is only a window of limited dimensions positioned inside the physical space of the viewer, the viewer is expected to concentrate completely on what she sees in this window, focusing her attention on the representation and disregarding the physical space outside” (Manovich 96). Thus, the screen exerts immense control over the viewer, and in turn filmmakers exert almost complete control over the environments within the window. This emphasizes the active role editors play in creating the story that exists within a highly selective screen. As the compilers of the images within the window, at every moment the editor is controlling what the viewers are looking at and where their attention is directed.

Although not the predominant means for creating perspective within cinema, mise-en-scene and sound contribute to the establishment of the viewer’s position in relation to the story. The effective utilization of sound within a film can act as a guide,
pointing us towards the relevant information on the screen: “Sound cues us to form expectations. If we hear a door creaking, we anticipate that someone has entered a room and that we will see that person in the next shot. But if the film draws upon conventions of the horror genre, the camera might stay on the man, staring fearfully. Sound can creatively cheat or redirect the viewer’s expectations,” (Bordwell and Thompson 292). Through the creative utilization of sound information can be disclosed or withheld, lending to the overall position viewers have when watching a film. Mise-en-scene can assist in developing the focus and attention within the filmic space, “much of the impact of an image comes from its manipulation of lighting. Lighter and darker areas within the frame help to create the overall composition of each shot and thus guide our attention to certain objects and actions. A brightly illuminated patch may draw our attention to a key gesture, while a shadow may conceal a detail or build up suspense about what may be present,” (Bordwell and Thompson 164). These decisions impact what viewers are observing and establish a certain mood or ambiance which directly shape the perspective or focalization. For example, in order to maintain suspense and horror, the viewer may be positioned farther from the happenings in the world through offscreen sounds and obscuring lighting to keep you knowing as little as possible until the proper climactic moments. These techniques cue viewers into the relevant information, establish the focalization, and position viewers within the world.

Despite the heavy authorial control in cinema, the viewer is still responsible for constructing the meaning in their mind, and of suspending their disbelief to follow the events and happenings in the story. Many of the roles delegated to the filmmakers in cinema, such as that of the editor and cinematographer, are shifted in virtual reality, and
the stylistic techniques used frequently in cinema, such as parallel editing, eyeline cutting, montage, and close up no longer translate in cinematic virtual reality.³

**Perspective in Cinematic Virtual Reality**

Cinematic virtual reality is exploring the seemingly paradoxical idea of expanding user freedom and authorial control simultaneously, resulting in renegotiations over who constructs perspective, how to buildup emotion, and maintaining story coherence. Many of the stylistic techniques and conventions of narrative and cinema which accomplish those tasks are being reshaped in virtual reality environments, where the emphasis is on 360° presentation of images presented with stereoscopic 3D and accompanying spatial 3D sound (Jaunt VR Studios). Cinematic VR constructs narrative stories where the viewer is afforded the ability to look around the space as the story is visually and audibly presented to them. Accompanying this experience is greater interactivity, through the user’s increased sense of agency accompanying the ability to control where they look within the environment. Further, there is a heightened sense of immersion, in that the screen is all encompassing. Returning to Marie-Laure Ryan’s model (fig. 1) juxtaposing the ability of various media objects to deliver immersion, interactivity, and design, we see the inherent sacrifices each category makes in differentiating itself. For example, while traditional narratives are ranked high in immersivity and design, they exhibit the lowest degree of interactivity. Contrastingly, role-playing games provide the greatest degree of interactivity, while potentially sacrificing design. When certain elements are prioritized—in the case of cinematic VR, immersivity and interactivity—subsequent sacrifices or

³ For more information on the plethora of techniques in cinematic editing and how they function to construct narratives in cinema consider *The Technique of Film and Video Editing* by Ken Dancyger.
trade-offs must be made as VR developers negotiate ways to best accommodate these goals. The result is new techniques, practices, and conventions.

As previously discussed, screens act as windows for the viewer to observe a carefully constructed spatial representation. However, VR complicates the notion of the screen as a window in that the screen now entirely encompasses the viewer. To quote Lev Manovich, “In contrast to cinema, where the mobile camera moves independently of the immobile spectator, now [in VR] the spectator actually has to move in physical space in order to experience movement in virtual space. It is as though the camera were mounted on the users head...The spectator is no longer chained, immobilized, anesthetized by the apparatus that serves her ready-made images; now she has to work, to speak, in order to see,” (Manovich 109). Whereas cinema relied on stationary viewing, virtual reality demands that power be granted to the viewer to choose where to look. When the screen is all-encompassing, visual strategies must also be adjusted: the result is fewer edits or cuts, more conscientious shooting environments, and the development of a new set of techniques--based on previous storytelling and cinematic conventions--which can be used to cue the viewer’s gaze toward what is most significant to look at. If a viewer is looking the wrong way, they may miss essential elements of the story, from subtle facial reactions to entire action sequences. Therefore, despite the elevated freedom to the viewer, even greater responsibility is placed on the filmmakers to guide perception in an intriguing and intuitive manner. Further, this relates to the previous notion of interactive narratives in that the viewer’s decisions impact the viewer’s understanding of the story unfolding.

In analyzing cinematic VR for its adherence to traditional cinematic form, we return to the metaphor of the viewer as the “camera eye.” Similar to cinema, cinematic
VR relies on the camera’s placement within the world to construct a perspective which is best suited to convey emotionality and progress in the story. One notable difference which extends the metaphor a step further is that in cinematic VR the viewer can choose where they look. The cinematic camera is entirely controlled by the filmmakers; the viewer’s agency to watch the scene is limited by this factor. However, now that the viewer can control the camera to a great degree in cinematic VR, the notion that the viewer is merely the camera eye is an incomplete comparison. Rather, I am making the argument that the role of the viewer in VR is shifting beyond the camera eye to encompass the roles of camera operator and video editor, as well. In this sense we become more than just the camera eye, and the act of looking is now also about editing and compiling a scene based on real-time choices made regarding where to turn your head and look. This is not to suggest that an editor within a virtual reality space is no longer necessary to or displaced by VR experiences, but rather, that the role of constructing perspective no longer rests solely on the filmmakers or VR designers. This also suggests that the stylistic techniques and system that is presented as the status quo in cinema must change to reflect this pivotal distinction in form.

Thus, the questions regarding the construction of perspective and techniques for progressing story and conveying emotion remains. If the viewer is granted the ability to look wherever they want, how do the creators make us look where they want? How do content creators ensure we focus on the details that are quintessential for understanding the meaning of the story? To what degree does our ability to forge our own perception of the events impact the delivery of the message? Stemming from the conventional tools of cinema solidified by Bordwell and Thompson, I offer a new set of techniques, or
classifications, which are pivotal in each cinematic VR object and which are inspired by the classical conventions of cinema. The four techniques I will expound include visual salience and selective lighting, spatialized audio, location and camera control, and editing. Each technique is reflected in various ways among each of the four media objects, the extent to which each object relies on each technique varies dependent on the story and goals of each.

To preface the proceeding sections, I will provide brief descriptions of each media object analyzed, all of which are cinematic virtual reality short films. This cataloguing is to ensure each cinematic VR object is given adequate introduction before diving into the methods each utilizes to construct perspective within their cinematic virtual worlds. Each of the case studies were selected for their ability—or lack thereof—to guide attention, position the viewer, and provide an immersive experience.

*The Great C:* This 37-minute short VR cinematic adapted from a 1953 short story of the same name. Released by Secret Location in October of 2018, the short depicts the story of a small community under the control of “the Great C,” a super-computer that demands a human sacrifice each year for appeasement. The camera is third person omniscient and follows the main characters as they embark on the journey to complete the sacrifice. The viewer is prompted to sit during the experience, and throughout the experience they can look wherever they would like in the frame. I selected this content for its ambitious translation of a written short story into a highly cinematic VR film. There are many cuts, fades, pans, and camera moves, with distinct scenes, many location changes and a standard 3 act structure. This project sought to translate as many elements
from cinema into VR as possible, making this object a unique experiment in adaptation and the primary reason for selecting this object.

*Awake Episode 1:* Awake is a 20-minute short cinematic created by production company Start VR. This film revolves around the psychological breakdown of Harry as he attempts to solve a mysterious lucid dream in which his wife is taken from him to another dimension of reality. This project was released in December of 2018 and utilizes a third-person limited camera wherein the viewer must prompt events to occur, and copious amounts of information are kept from you. In this object the role of the viewer is as observer and detective, and you can walk around the space to get closer to clues and read notes. Further, you are required to interact with objects in the world to progress the story forward. The story requires you to complete actions with your controllers--failure to do so will cause the story to stall. I selected this short film for its interactive features, ability to move around the space, and methodical use of editing.

*Pearl:* Pearl is a Google Spotlight story with a run-time of 6 minutes created in 2016 and was the first VR film to be nominated for an Academy Award. The story is about a girl and her father, and their relationship over time as they travel across the country in a beat-up car that serves as their home. Third person omniscient perspective is utilized, with a heavy emphasis on music which guides and progresses this montage-esque story. Here, the viewer can look wherever they want within the frame, though their position within the world never changes. The camera is placed on the passenger seat of the car and does not move. I chose Pearl because although there are frequent cuts and large time progresses, they become invisible through the methodical use of music and sound.
Tree: Tree is a VR short initially released to film festivals in 2016 and added to Viveport in January 2019. This VR short film utilizes first-person perspective to place the viewer in the position of a growing tree, where you experience the life of a tree in the rainforest. *Tree* was created by the New Reality Company to raise awareness of deforestation. In this short the viewer must stand, and as the tree grows the viewer can move around the space, thus causing the tree to shake. Additionally, your arms correlate to the branches on the tree and you can shake them around. I decided to include *Tree* in this paper because of the filmmaker’s attempt to present a narrative in first-person that is unlike other cinematic experiences. The editing is minimal, as is the story, however they utilize several techniques which serve to continually shape perspective and grant the use as much freedom as possible.

Technique 1: Visual Salience and Selective Lighting

Visual salience is a critical factor which shifts attention and drives immersion. To quote Jason Jerald, “Salience is the property of a stimulus that causes it to stick out from its neighbors and grab one’s attention. The attentional capture reflex occurs due to the brain having a need to update as quickly as possible its mental model of the world that has been violated” (Jerald from Coren et al., 1999, 149). There are many ways in which this can be achieved; using color, lighting, direction, shape, and size various objects can be separated from their surrounding environment. The purpose of visual saliency is to direct attention to the most relevant information in an environment. Stop signs are red so that they stand out from the surrounding environment and we notice them, poisonous frogs are brightly colored, so animals know to stay away from them. Visual salience is a
tool of mise-en-scene, where in filmmakers can intentionally foreground elements which are most significant. Just as mise-en-scene in film guides and immerses, in cinematic VR mise-en-scene acts to guide the viewer and establish relevance in the 360° environment. While other aspects of mise-en-scene such as the acting and color composition still serve to enhance the cinematic world, the primary focus in this paper as observed through the media objects is the use of visual salience and selective lighting in directing attention, creating perspective and driving immersion.

The concept of visual salience allows for viewers to look wherever they chose, while simultaneously granting the creator authority to emphasize what is relevant. Each media object utilizes visual salience to varying degrees, with the most apparent method in cinematic VR being the use of selected lighting. With selective lighting certain locations, objects, and actions are spotlighted to draw attention towards and provide immersion in the world.

The film *Awake* relies heavily on the idea of visual salience to progress the story, construct the viewer’s perspective, and spark interaction within the environment. Using selective lighting, color, and most importantly, direct prompts which the viewer must complete to progress the story (fig. 4), this film maintains both immersion and interactivity. In flashbacks the lighting and colors are spotlighted so that only the area where the actors are interacting is detailed. While this may be due to a lack of time or resources, the effect is that viewers are directed to attend to the small area that they can see in full detail. Further, there are specific items in the world that play a significant role in the events of the story, such as a teacup and cube, which are noticeably different from their surrounding environment. Both are brightly lit, causing them to stand out from the
surrounding dark, bleak, and cold environment. Salience cues the viewer into these items and signifies their importance to solving the mystery.

There are several times in the story experience where the viewer is prompted to complete an action in order to progress the story. A small globe which is clearly distinct from the surrounding environment appears and the controllers vibrate (fig. 4). Inside the globe a projection of another space can be seen, such as in the figure below with the image of the telephone. When viewers tap the globe, they are taken to the new space within the room, and the story progresses. This is reminiscent of ergodic literature, wherein nontrivial effort is required to progress the story, signifying that cinematic VR is going in the direction of greater interactivity and user involvement. The technique not only provides an opportunity for the user to feel as though they are interacting in the space, but also places the viewer within the world of the story and constructs the viewer’s perspective as critical in the story progression.

Fig. 4. The image on the left shows the globe that the player must touch and interact with. Once you successfully touch the globe, the viewer is taken to a new location in the room, reflected in the right image. Viewers are then instructed to pick up the phone themselves as noted through the glowing light above the phone. Taylor, Martin, director. Awake Episode 1. Start VR, 2018, starvr.co/project/awake/.
In *The Great C*, light is used as a motif throughout the story to maintain coherence across the multitude of locations and environments. The protagonist is sent on a journey to sacrifice himself to the supercomputer and must follow light posts to get there. Thus, a long network of lights spanning the entire journey becomes the way the creators orient the viewer’s perspective and maintain clarity (see fig. 5). Because of the lights, they are aware of their placement in the environment, providing a quick way to orient themselves with each cut. High degrees of visual salience are used in several montage sequences to guide perspective and present information which is critical for understanding the story. In these sequences, the background is black, and the foreground is lit up, often revealing specific details which home in on the emotional and compelling elements of the story. The viewer doesn’t have to focus on the entire environment, which can be an incredibly fatiguing experience, and instead can focus on the ideas being communicated. This takes on a similar feel to video game cutscenes, wherein temporary control is taken away from the viewer so that the relevant details of the story can be revealed before you can continue gameplay. Here however the viewer still can look wherever, but the saliency of these scenes is so strong that if they look anywhere besides immediately ahead of them, they will see nothing but black.
In *Pearl*, because we are inside the car for the entire duration of the film, the primary creative goal is to maintain interest with the actions occurring inside the vehicle. The lighting is realistic and does not attempt to single out any specific elements or characters. There are times when lighting shifts, rather than lighting the interior of the car the creators light the external environment. This occurs when the characters are outside the car and the focus is not inside the vehicle, as seen in figure 6. Whereas the use of lighting is subtle in *Pearl*, the use of color to convey the emotionality of the story is apparent and critical to the mood. Bright colors and skies are keynote of the scenes with happy memories, meanwhile more negative memories are noted by darker colors and nighttime environments. Even the protagonist’s wardrobe changes to reflect her adolescent delinquency. The subtle use of salience within this short film provides greater freedom in that you feel present in the vehicle, however, there are no compelling visual elements to make any aspects stand out. Your perspective as a passenger along for the

---

*Fig. 5.* In the left image, the protagonist is framed by light in a wide shot. Following that shot is a cut to the image on the right, where the protagonist is still lit by the same light. Here, light serves to orient the viewer within the world. The rest of the environment is very dark, but the light illuminates the path forward. Miller, Steve, director. *The Great C.* Secret Location, 2018, secretlocation.com/experiences/the-great-c.
ride is established through the minimal salience. As a passenger, you chose where you look in the car. Other elements work together to compel the viewer to look where the story is happening and not out the window.

Similarly, in *Tree* there is a lack of immense visual salience compelling the viewer to look a certain direction or focus on details. This is on par with the story; you become the tree and therefore the responsibility is on you to explore the environment. Eventually your tree grows branches which correspond to your arm movement, as you grow you experience new surroundings which serve to immerse the viewer and provide a realistic experience. Lighting plays a significant role in the beginning and at the end of the film, and in each instance is used specifically to increase the empathetic responses. In the beginning, light shines through the dirt and your seed grows into a sapling which

*Figure 6. The perspective and camera position do not change, however the lighting shifts to place emphasis on the actions. In the left image, the background is over-exposed and blown out to maintain our interest inside the car. In the right image, the characters have left the interior of the vehicle, and now the interior is dark while the exterior is properly exposed. The lighting makes certain areas of the frame more salient, and noticeable, than others. Osborne, Patrick, director. *Pearl. Google Spotlight Stories, 2016, atap.google.com/spotlight-stories/*.*
bursts through the Earth. Here, light serves to direct your attention upward, and to create a more realistic and immersive environment so you can imagine yourself as a budding tree emerging into the world (fig. 7). Similarly, at the end the lighting is highly salience, as the dark environment juxtaposed with the light from fires spreading all around you is used to call attention to the helplessness of your perspective as a tree (fig. 9). You can’t run, and the creators calling attention to the encompassing fires by darkening the background while the remaining forestry burns drives the message home.

The use of visual salience is an extension of that literary theory of focalization, wherein certain elements of the story are foregrounded to indicate relevance. In cinematic virtual reality, selective lighting and visual salience accomplish the same goal. When creators choose to emphasize and call attention to certain elements of the virtual environment, they are creating a specific mood, and disclosing certain information in a particular way to the viewer.

![Fig. 7. Light serves to direct the viewers attention upward in the left image, as you are a budding sapling emerging from the dark ground. Further, the use of light to distinguish your tree arms from the background environment emphasizes the focus of the story, that you are the tree. Zec, Milica and Winslow Porter, directors. Tree. New Reality Company, 2016, www.treeofficial.com/](image)
Similarly to focalization in literature, this idea of visual salience provides the authors an opportunity to shape the viewer’s perception of events. Further, this idea relates to the concept of editing as a tool for conveying emotional emphasis. In editing, shots are compiled to elicit a specific emotional reaction. Here, certain information is presented to elicit a specific emotional reaction and establish an aesthetic. To advance the suspense and horror, *The Great C* purposely lights certain scenes darkly to heighten the drama, whereas *Tree* attempts to maintain consistent lighting to make the environment as realistic as possible.

**Technique 2: Spatialized Audio**

Virtual Reality as an immersive and interactive experience is as dependent on the construction of sound in a 3D environment as the presentation of sequential images, if not more so. Returning to Jerald: “Spatialized audio is sound that is perceived to come from some location in 3D space” (Jerald 240). The use of spatialized audio is a major tool used to direct us where to look. VR’s presentation of spatialized audio provides creators enhanced control and the ability to guide perspective and direct attention in this cinematic content. Audio constructs the environment and immerses the viewer by surrounding them with sound parallels the ideas of cutting for emotion and using camera placement as tools for immersion seen in cinematic constructions of perspective. The foregrounding of certain sounds to direct attention and cue the viewer is a dramatic tool as much as it is a realistic one. Not only does sound exist within VR to accurately construct the environment, but to create emphasis and guide perspective. Further, your placement
within the VR environment impacts what you hear; rather than only the camera dictating the environment and the world surrounding the camera, the audio too surrounds. Each virtual reality object relies on spatial audio to varying extents, as a means for capturing and guiding attention. The main difference occurs in the creative use of sound to progress the story and heighten the emotional impact of the story.

The most audio dependent short is *Pearl*, whose story revolves around music. What begins as diegetic music, Pearl hitting play on an old cassette tape of her father singing, prompts the beginning of the montage in which her father performs the song. Halfway through the short the song transforms into a non-diegetic element which carries through the story of Pearl and her father’s relationship. The song is titled, “No Wrong Way Home,” and was created by Alexis Harte for this film. Not only does the song provide coherence and hide the frequent cutting but provides a unique and emotional touch to a heartwarming story. The song feels personal and we are invited into the space that the characters call home. As inherited from continuity editing in cinema, most of the cuts occur on action, and are in perfect sync with the song, further minimizing feelings of overt shifts in time or space. For *Pearl*, the ability to spatially surround the viewer with highly emotional audio directs attention seamlessly.

*The Great C* relies heavily on audio to construct the world and help the viewer determine what is worth looking at. Because the film is loaded with visual information, and the constant cutting and camera movements can be jarring, audio serves as an indicator of where to focus. The film cuts from close ups to medium shots and into wider shots, all while dialogue and action are occurring. Here, the audio bridges the gaps between each cut and maintains a sense of orientation within the environment. The audio
plays a significant role in distracting attention away from and subverting the presence of aspects the viewer cannot control, such as the frequent editing or location changes. The nondiegetic music additionally serves to heighten the emotionality of the story, despite the lack of interactivity and control there remains a compelling and engaging message.

In *Awake*, audio is used in specific instances to call attention to Harry’s frame of mind and emotional state. The telephone ringing, the metronome ceasing, the bird beginning to screech are all audio cues which serve the purpose of cueing your attention to a heightened emotional experience, thus adding depth to the room. Once the bird ceases flailing and dies, a psychological trip occurs and Harry nosedives into a traumatic memory. The audio contributes to the surrealist and psychologically dense aesthetic of the film, serving to disorient and confound the audience at times. In montages we hear Rose’s voice as though she is an omniscient being in the world, her nonsensical sentences add to the confusion and we begin to understand why Harry has lost his mind. Here, audio serves the goal of distracting attention away from the editing as well. Rather than focusing on the fade outs, our attention remains fixed on the confusing omniscient voice of the protagonist’s love interest or creepy instructions of the omniscient antagonist.

*Tree* relies on the use of spatialized audio to guide attention to a lesser extent than the other objects. The high degree of freedom granted to the viewer to look around and observe, matched with a minimal and slowly progressing plot, yields a lack of necessity to guide attention and maintain focus on any part of the environment. The experience is meant to be personal, and the audio primarily serves to immerse the viewer in the rainforest environment. The main use of audio to cue attention comes when the forest fire begins. Shrills of wildlife fleeing and dying in the distance prompts the viewer to look in
the specific direction they were heard, and upon looking the fire is revealed. Your
attention is guided using audio and your perspective shifts as a result of the emotional
and empathetic impact of the screaming animals. At this moment the scene cuts, and you
are no longer observing the world from the perspective of the tree. You return to a
position of omniscient observant at the end of the story; however, the jarring audio and
visuals bring about a new perspective on deforestation.

**Technique 3: Location and Camera Control**

To maintain creative control, whether for the intent of preserving the story,
successfully immersing the viewer, or any other specific rationale, creators rely on
constructing bounded spaces for the viewer to interact with. This is primarily achieved
through location control and camera placement within the environment. While the latter
is specifically an issue of cinematography, I have chosen to place location control here
instead of in mise-en-scene because of the relationship location and camera have with
one another in cinematic VR. Because the screen undergoes significant changes in VR,
the camera must adapt to meet those demands. The camera in film maintains a great
degree of flexibility in that the crew can pick it up, move it around, and plant it wherever.
To change the camera’s direction all that the crew must do is shuffle around so they are
not in frame. The new VR camera is recording in all directions, no longer can the crew be
situated conveniently out of frame, because the frame is now encompassing the camera.
Because of this significant change, location becomes a critical concern for the
cinematographer. How the cinematic VR cinematographer, camera crew, and director
control and manipulate the camera depends on how they can interact with the location and how the camera is placed within the environment.

In this discussion, the negotiation between user freedom and creator control becomes most apparent and the conventions of cinema break down. Whether creators choose to limit or expand their control of camera movement and location control plays a significant role in how perspective is constructed, the extent to which attention must be directed, and the role that editing plays in the story construction. In this domain the case studies are varied in their approaches to constructing their respective stories.

*The Great C* spans over many locations, and has incredibly ambitious camera movements that involve panning, tilting, and large sweeps across expanses of space (fig. 8) Viewers still choose where they would like to look within the environment, however the frequent cuts, camera movements, and location changes can be a disorienting and fatiguing experience. Due to fatigue and information overload viewers surrender to viewing what is automatically presented in front of them, and in this regard minimal effort is required to direct attention. The extravagant camera movements discourage frequent head movement, resulting in a VR experience that makes few concessions to the perspectival agency afforded by VR. Said another way: in deciding to expand the use of locations and camera movements, user freedom suffers dramatically in *The Great C*. The creators exert a great deal of control over the cinematography and editing, which become so locked down and tied to the rhythms of conventionally edited cinema that the viewer can’t really experience what VR sets out to provide.
The Google Spotlight cinematic VR story *Pearl* technically has many location changes, as you are travelling in a car across the country, however the camera stays in one place throughout the duration of the film. As a result, the perspective remains constant; you are planted in the passenger seat of the car the entire time as an observant along for the ride. Here, the creators chose to surrender their control over the camera for the ability to change locations and present a montage. This grants the user freedom to look around the environment without feeling fatigued, however it does not ensure that a viewer’s attention will be maintained during the duration of the film. Viewers can easily choose to look out the car window the entire duration of the film and are provided no visually compelling reason to do otherwise. The main force which drives attention in this short film is the music and audio. This is one sacrifice VR creators must risk when granting viewers with increased freedom, however, to most viewers that want to be
immersed the increased interactivity leads to greater immersion. The ability to look around while inside the car is meant to make you feel as though you are a passenger along for the ride of this heart-warming story.

*Tree* occurs in one location, the rainforests of South America, and throughout the duration of the film the camera is steadily raised upward as your tree grows. As the camera is elevated your tree enters different biomes within the rainforest, making it appear as though you are changing environments. The camera is manipulated along the x-axis and the while the viewer has complete control over where they chose to look, they cannot control the rate at which the camera is raised. Similarly to Pearl, the creators decision to minimize their control over the camera yields greater opportunity for viewer freedom, and can lead to a more immersive experience. The perspective is understood almost immediately, you are in the first-person experiencing the life of a tree. Only at the end, when the fire begins to consume the tree we were embodying, is there a cut to a wide shot overlooking the entire forest on fire (fig. 9). The camera shift represents the end of the tree’s life; at that point viewers are encouraged to reflect inwards on their experience. The creators scarcely impede or exert control, yielding to an experience lacking guidance or cues to alert attention. The viewer is encouraged to look wherever they want and may zone out or lose attention altogether as the story lacks strong attention cues through the surrendering of location and camera control by the creators.
Awake takes place in a single location, the study room in the main protagonist’s house. However, the camera’s position is changed several times throughout the film, though no pans, tilts, or raises occur. The only manipulation of the camera that occurs is its placement within the room. The viewer is given control over where they look and can even move around the space and get closer to elements such as the fireplace, chairs, or books during the experience. Because the location does not change, the creators can get away with changing the camera’s position throughout the film. The viewer does not need to spend a great amount of mental energy calibrating their new position within the environment, which also serves as a tool for masking edits (fig. 10). Despite the control over the location, this film lacks a clear perspective through which the experience is viewed. At times we are observing from a third-person omniscient perspective, other times the camera is placed over or on top of Harry’s shoulder, suggesting a first-person

Fig. 9. The image on the left is the last frame before there is a fade to black. Viewers see their own tree that they’ve been embodying light up in flames. The image on the right is the next shot, and it is of the same tree up in flames. Zec, Milica and Winslow Porter, directors. Tree. New Reality Company, 2016, www.treeofficial.com/.
viewing. Awake presents the greatest opportunity for interactivity through the globes with which viewers must touch to progress the story. This tool is made possible through the limited location and positions within the environment that the viewers can potentially warp to. Here, the increased limitations on location results in greater freedom for the viewer to explore.

![Fig. 10. The left image is a wide shot of the room. The globe icon in the center of the area is a prompt and touching it with both controllers causes the scene to cut into the next shot, which is the image on the right. Because of the size of the space and direct cues, the cutting is not disorienting, and the immersion is maintained. Martin, director. Awake Episode 1. Start VR, 2018, startvr.co/project/awake/.
](image)

The idea of increasing authorial control over location and camera as critical for increased immersion and interactivity seems counterintuitive. However, limiting the space within which users can interact may yield greater immersion in that viewers are not overwhelmed with seemingly limitless options and arbitrary destinations. Increased authorial control works hand in hand with increased user agency, if authorial control is increased with no subsequent trade-offs or increased user agency, this does not synergize. This is reminiscent of the idea of Alberti’s window in that the limited screen, or window
into another space, allows the creator the ability to shape perspective and yield an impressive visual feat. Further, through the use of the term cinematic in the category of cinematic virtual reality suggests a particular type of relationship between the viewer and creator. Viewers are ready to submit their unfettered freedom to be immersed in a compelling story, and therefore a greater degree of authorial control is required in cinematic virtual reality as compared to more game-based virtual reality.

**Technique 4: Editing**

The role of editing in narrative-driven, Hollywood-style cinema is to be as invisible as possible, and in this regard editing in cinematic VR shares the same goal. If cuts are too noticeable in virtual reality, viewers may be taken out of their immersed state, or worst of all, may become physically sick. While Walter Murch’s rule on cutting first for emotion remains relevant in virtual reality, we can no longer assume the viewer will even be looking where they “ought” to be when the cut is made. Even though the viewer has adopted responsibilities previously exclusive to the camera operator, cinematographer, and editor, cinematic VR developers also have new and expanded responsibilities in relation to these roles. The conscious decisions of cinematic VR creatives to use specific stylistic techniques impact the extent to which other techniques play a role in the construction of the story. Here this is made most apparent, as the editing techniques adopted in each short film are a consequence of the stylistic decisions made previously in the production process.

In order to accommodate the excess of locations and camera movements in *The Great C*, editing plays a significant role in constructing the story and perspective. With the frequent change in camera placement and location editing becomes difficult to hide,
as frequent cutting is required to progress the story forward. There are many fades to black, cutting on action, and cutting on visual effects, which represent the filmmakers attempts to ease the viewer into the cuts. Through the use of the lighting motif in mise-en-scene, the filmmakers are able to ground viewers to an extent and orient them in the midst of frequent cutting, however the editing only works when the viewer is significantly fatigued and discouraged from looking anywhere besides where the camera automatically directs them. While the experience and story are interesting, the overwhelming control the filmmakers exert over the camera yields a need for invasive and frequent editing, which has a significant impact on the processes of immersion and focalization. Additionally, rather than utilizing editing as a tool to enhance emotion, it becomes a tool for progressing the story forward and compiling a bunch of cool shots.

In *Pearl*, due to the stationary camera the editing is easily obscured, as there are no jarring changes in camera placement between cuts to take the viewer out of their immersed state. When cutting does occur minimal work from the viewer is required to reorient within the environment as they understand quickly that their position is not changing. Further, the emphasis on the musical aspect of the short film acts as a distraction from the frequent cutting by providing overarching coherence. Nearly all the cutting is done quickly on the beat, as though you are blinking your eyes, which yields itself to a more immersive experience. Here, despite the excessive use of cutting to progress this montage which spans over two decades of the characters’ lives, the other elements work together in harmony to hide the edits and bring out the highly emotional message presented.
The slow and steady rise of the camera accomplishes the tasks of editing in *Tree*, over time we are privy to new locations and environments without needing to cut. The location and mise-en-scene play a significant role in determining the minimal responsibilities of editing within this short. Because we are placed in the perspective of the tree, we are grounded to a single location and are encouraged to identify with that tree. Any editing would only serve to take us out of that intimate relationship. There is one fade to black towards the end of the film which occurs when the film shifts the viewer’s perspective from that of the tree to an omniscient observer of the forest. Even here, the cut is natural and motivated by emotional reasons which makes the edit nearly invisible. In this case, editing was sacrificed for the sake of a highly immersive and relatable first-person experience.

*Awake* cleverly utilizes editing to heighten the audience’s immersion into the main protagonist’s psychological state of being, at times at the cost of presenting a coherent storyline. Accompanying the jumps in space around the room, are jumps in time that can be confusing and difficult to comprehend. They occur as Harry enters a dream state, which signifies a disorientation with a linear time frame and a merging of reality with fantasy. Almost all the cuts are fades to black and done slowly, perhaps to provide viewer’s adequate time to properly reorient themselves in the environment. Regardless, the cuts in time and space lead audiences to question what elements are most significant and from whose perspective we are receiving information. However, this might be an intentional decision by the filmmakers to reflect Harry’s inner state of turmoil and psychosis, as well as to heighten the emotional impact of the story and connect the viewer with the character. Additionally, cuts will occur only after prompted by the user’s input,
again returning to the globes within the world that the user interacts with. Here, though the edit is calling attention to itself, the increased role that the viewer plays is enough of a distraction.

In each case study, the role that editing plays is significantly impacted by the techniques utilized in the production process. For example, whereas in *Tree* the creator’s intentionally limit the location and perspective, in *The Great C* the creators strive to utilize numerous locations and camera movements. As a result, the editor can play several different roles in cinematic VR, from an absent figure to the main individual responsible for piecing together the story. Clearly, editing is not going anywhere in cinematic VR, however due to changing stylistic techniques and enhanced viewer control, the role of the editor is altering to accommodate the demands. Whereas in cinema, the editor plays a distinct and critical role regardless of the other stylistic techniques, in cinematic VR due to the plethora of factors impacting perspective (mise-en-scene, cinematography, user freedom) the editor’s title is not as obvious.

**Analysis**

Currently, research into VR is attempting to formulate a taxonomy by which we can label and quantify how attention is guided in virtual reality environments. In their article titled “A Taxonomy for Deploying Redirection Techniques in Immersive Virtual Environments,” researchers differentiate the ways in which reorientation and repositioning occurs. They propose that both reorientation and repositioning can either be overt or subtle, and continuous or discrete, allowing for four permutations each (fig. 1). To summarize their findings, the researchers concluded that audiences prefer subtle, rather than overt, redirection and reorientation (Suma et. al). Overt attempts to shape
perspective are perceived as jarring and counter to immersion and should therefore be avoided. Their findings are relevant for understanding how viewers perceive VR environments, however this taxonomy is focused more on the classifications of attention, rather than contextualizing techniques for guiding attention.

Another study conducted by researchers at the University of Copenhagen sought to provide a taxonomy for guiding user’s attention specifically in cinematic VR. Their taxonomy is similar to that of Suma et. al., however, their parameters are slightly different (fig. 1). Though not statistically significant, there research seemed to suggest that, “forced rotation may hamper the sense of presence,” (Neilsen et. al. 231). Both findings serve to bolster the taxonomy of cinematic techniques I develop by suggesting that certain characteristics or cues work better than others, therefore impacting the techniques utilized and the degree to which creators can exert authorial control over the environment. They both note that their goal is primarily to inform further studies into exploration of the effectiveness of various techniques for guiding attention and redirecting in VR. In their closing remarks Neilsen et. al. remark: “Besides from providing an overview of the space of potential approaches, the taxonomy may also inform future studies exploring the effectiveness and naturalness of varying cues,” (232). They nod towards opportunity for growth and development, and in this regard the quantitative data collected provides credibility to the quantitative findings I observed. In Nielsen’s findings, they speculate that nondiegetic cues are generally more detrimental to feelings of presence. This indicates that techniques such as visual salience or spatialized audio that is diegetic will yield a greater sense of immersion. Because research seems to
suggest that users want less overt cues, developing techniques to build the world will be essential in guiding users without being obvious.

When determining the responsibilities of the viewer and creator within virtual reality the most critical goal is achieving balance. Finding the middle ground between wandering aimlessly in an unexplored forest and hiking a well-worn trail is an apt analogy for shaping viewer agency in or interaction with cinematic VR. Whereas viewers want to have the ability to roam and construct the story as they please, a complete lack of boundaries and parameters yields a meaningless experience. On the other extreme, an experience that is too rigidly constructed and bounded lacks immersion and interactivity, rendering the viewer a slave to the story.

![Diagram of Taxonomy of Redirection Techniques](image)

*Fig. 11.* The chart on the left is from Suma et. al. and presents a guideline for supporting natural redirection and reorientation in VR. The chart on the right is from Neilsen et. al., and offers a taxonomy which categorizes the types of cues for guiding user attention.

Sometimes limitations hinder creative growth, however if we look to 3D cinema, we can observe an expansion of creative capabilities in a time of technical limitation. Due to the heightened amount of sensory information from the addition of a Z-space, or third dimension, audiences require more time and brainpower to comprehend the moving
images presented to them. This places several restrictions on the cinema editor, quoting visual effects artist Bernard Mendiburu in his book *3D Movie Making*:

There has been a long-lasting consensus that 3D should be cut its own way. *Which* way was not really defined beyond being at a slower pace than 2D, considering that 3D images are more complex to process visually. Not only is 3D reading time longer, but the audience tends to scan the whole scene before going back to the subject. The detractors of such an approach believe that the technique should not influence the art, that it’s the DP’s job to make changes that match the script’s intended rhythm. We see here that this debate on editing style was most likely decided one way or another when the picture was shot. You will either adapt the edit to the depth or adapt the depth to the edit. (151)

This raises a concern similar to that in cinematic VR: does editing make an attempt to define its own unique way of editing in VR, or does it adapt to meet the demands of the cinematography and mise-en-scene? Regardless of the answer, if done properly, a 3D film can provide an even more immersive experience than 2D cinema, indicating the potential associated with successful negotiation.

The methods by which attention is maintained and perspective constructed adapts to meet the demands of the medium. In VR, new and previous tools must work together to accomplish the goal of presenting a fantastic story experience. The cinematic VR objects that attempt to stray from film and branch out are utilizing the cinematic conventions in more abstract ways as compared to the shorts which are intentionally attempting to translate as many aspects from cinema as they can. The content creators of *The Great C* state their objective in an interview with Stephen Reid, “We made it a goal for our version of *The Great C* to ‘feel’ like a cinematic experience. We focused on adapting film language and techniques, like editing, composition, scene transitions, camera movement, and pacing for VR, to give the piece the tempo and excitement of a
big screen movie,” (Reid “Viveport Q&A”). Contrastingly, the creators of _Tree_ had a different objective in mind, “For _Tree_, one of our main goals was to place users into the ‘shoes’ of nature. We made our users the center of the _Tree_ experience; each one conducts their own story, so the piece becomes totally unique to each person. _Tree_ is quite ambitious from a technical standpoint. The success of the piece relies upon making users feel like they’re consistently and steadily growing as a Kapok tree, while also surrounded by dynamic lighting and shadows generating in real-time,” (Reid “VR for Impact”). As a result of different goals, the stylistic techniques were utilized differently.

As a result of this research I propose a new stylistic system and techniques by which cinematic VR progresses its formal narrative system, leading to the classification of cinematic VR as a new form. The four techniques proposed, visual salience and selective lighting, spatialized audio, location and camera control, and editing, do not exhaust the ways content creators are exerting or surrendering authorial control but reflect the critical ways perspective is constructed within cinematic VR. As with the cinematic stylistic system (cinematography, mise-en-scene, sound, and editing), different filmmakers choose the degree to which they use each technique and in what ways. However, the extent to which every cinematic VR short utilizes each tool, matched with an analysis of their commercial success can work to establish solid criteria by which we can standardize cinematic virtual reality.

A primary change from film to cinematic virtual reality is in the role the viewer plays as a camera operator and video editor when inside a cinematic VR experience. As a result of this renegotiation adjustments must be made by the cinematographer and editor in particular to adapt to this shift in responsibility. This shift should come naturally for
viewers, as we are already editors in our own lives. Returning to Walter Murch’s book on
editing titled *In The Blink of An Eye*, he states that in our everyday experiences blinking
serves as a sort of cutting and filtering information, “… our rate of blinking is somehow
gearied more to our emotional state and to the nature and frequency of our thoughts than
to the atmospheric environment we happen to find ourselves in. Even if there is no head
movement, the blink is either *something that helps an internal separation of thought to
take place*, or it is *an involuntary reflex accompanying the mental separation that is
taking place anyway*” (Murch, 62). With this in mind, we see that the ability to choose
where we look and for how long in virtual reality mimics that procedure of information
processing that we naturally conduct all the time. All that is required is presenting a
compelling message worth focusing on and our brains automatically process the
information and guide our attention.

If users want enhanced freedom in cinematic virtual reality, filmmakers must find
reliable means to tell a story which accomplishes the demands of the audiences. Lacking
reliable techniques by which stories can be constructed makes the process of sparking
mainstream interest more difficult and unstable. Because VR is an expensive endeavor,
creators must find ways to ensure the experience is not only enjoyable, but mesmerizing.
Compelled viewers will adopt the role of editor seamlessly, as natural curiosity and
enthrallment in a great story will overcome any desire to “break” or subvert the cinematic
VR experience with counter-productive looking. The four techniques proposed are vital
to reconciling VR conception with traditional narrative and cinematic storytelling to
create successful cinematic VR experiences. These techniques work together to create
and maintain a coherent viewing perspective despite the changing notion of the screen
and ability of the viewer to look around wherever they want in the environment. The result is a synergistic relationship, wherein each technique bolsters the others and truly elevates the form to an art.
CHAPTER 5
CONCLUSION AND OUTLOOK

Reflection on the Process

Understanding the development of media and the ways in which each medium adapts to best present the story and message is as much an issue of theory as observing practice. With the current state of cinematic VR as a budding industry, minimal theoretical work for understanding narrative and aesthetic standards in VR has been conducted. This lead my research down the paths of narrative and cinema to contextualize what I was observing within virtual reality that has not yet been concretized in any academic literature. Because of this, I also had to immerse myself in cinematic virtual reality and reflect on what creators are currently doing in the industry and the ways in which they are attempting to solve issues of viewer agency, editing demands, and authorial control.

During the research and writing process my views on editing and the debate between authorial control versus user freedom within virtual reality shifted. Initially, I believed that I would see complete erasure of editing in virtual reality, with the thought that any cuts made would surely take the viewer out of the immersive experience. Rather, ideas such as cutting for emotion and story over continuity, cutting for dramatic emphasis, the use of montage to span over great lengths of time were translatable to cinematic VR in a way that encouraged immersion and emotional connection, rather than hindered. The key struggle that arose instead was that of determining when the editor was
responsible for cueing the viewer and when to surrender control to the viewer. Further, I came to realize that my initial thought that users would demand freedom and expect all virtual reality content to be purely ergodic was incorrect. Viewers enjoy surrendering freedom and work for the sake of being immersed in a meaningful story that has a strong emotional impact. As with cinema, viewers are willing to let creators take the reins. The key is finding a balance between the user's desire to passively experience content and to have an interaction with what they are viewing.

While attempts to perfectly translate cinematic conventions into virtual reality prove to be messy, many of the goals for each medium are the same. Cinema and cinematic VR strive to immerse viewers, tell a great story, and make an impact on the emotional state of viewers. However, the methods by which these goals are achieved changes depending on the goal. Additionally, I learned that the degree of interactivity or immersion is not the only nor most significant criteria for determining the value of a cinematic VR object. While one goal of cinematic VR may be to allow viewers the ability to have greater interaction within the story environment, currently the more successful experiments are those that maintain a significant degree of authorial control. This seems to suggest a relationship between the granting of greater user freedom and subsequent increase of creator authority to meet those demands.

Areas for further inquiry

As Cinematic Virtual Reality continues to grow, develop and experiment, the relationship between viewer and creator will continue to negotiate. Currently, cinematic VR encompasses works which follow narrative cause-and-effect over space and time wherein the viewer does not have the ability to significantly alter or modify the sequence.
of events. As our technology improves and computer become more powerful, perhaps with the advancement of artificial intelligence, viewers will be able to consume and engage with narratives in a manner that alters the story’s course of events and has even greater degrees of freedom. The best example to point to is Interactive Theater, where the fourth wall is broken in that audiences interact with the performers to progress the story and make decisions about how the world and story is going to unfold, and what information is revealed. Of course, this works with theater because the agents reacting to audience input are human and can adapt to moderate and steer the story in the right direction, but how would computers intuitively perform the same tasks?

Additionally, there are severe limitations on location, camera, and editing within virtual reality that must be overcome in order to present a world design and interactivity that matches the level of immersion offered. The development of solid criteria will come over time and with experimentation, and through exploration into live theater, ergodic literature, and eccentric cinema, answers to these pressing questions regarding agency and perspective may be found. The short films analyzed in this thesis do not present multiple perspectives, as cinema and literature do, and in this regard cinematic virtual reality has boundless potential for growth. Nor do they incorporate game into their design, which is another growing trend in VR, which has even greater potential in redefining the relationship viewers have to objects in a VR environment.

In summation, while cinematic VR has strides to make in developing its form and utilizing techniques in an effective and meaningful strategy, the potential is boundless. The main limitation is VR’s ability to tell encapsulating and awe-inspiring stories that compel viewers to come back again and again. Visual effects and impressive
environments may be the hook that catches people and intrigues them into trying VR, but the story is what keeps them interested and invested.
REFERENCES

Aarseth, Espen J. Cyber


