VOLUMETRIC VIDEO AND THE FUTURE OF VIRTUAL REALITY

SLAVICA CEPERKOVIC

A DISSERTATION SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CINEMA AND MEDIA STUDIES

YORK UNIVERSITY

TORONTO, ONTARIO

August 2022

ABSTRACT

This multi-modal dissertation focuses on research creation of haptic cinema and new ways of designing immersive story experiences with volumetric video. Recent technical innovation of volumetric video allows new definitions of haptic cinema to include self-gaze by artists to be explored in an accessible way. The research creation developed were prototypes using volumetric video in virtual and extended reality (XR) environments. Various tools were considered during the research creation process including Unity, Unreal Engine, Depthkit, Mozilla Spoke Hubs, Ableton, and Radical.

ACKNOWLEDGMENTS

I would like to acknowledge and thank my supervisor Dr. Caitlin Fisher and committee members Dr. Sharon Hayashi and Dr. Graham Wakefield. Your courses inspired and helped me build a strong foundation for this applied research. I would also like to thank graduate program assistant, Kuowei Lee, and graduate program director, Dr. Michael Zryd, for your support and dedication to graduate students.

TABLE OF CONTENTS

LIST OF IMAGES

PREFACE

This document frames the applied research process with a multi-modal approach. The introduction guides the reader through the research questions and considerations. It also articulates the questions derived from the applied research creation process and later used as a voiceover in the VR work. It is suggested to the reader to start by reading the introduction and then referring to the appendix and watching the video walkthrough ***"****Documentation – Considerations for Volumetric Video in VR"*. This is a walkthrough of the virtual reality research and articulates in a concise way what is illustrated throughout the paper and will provide additional context to the reading.

The context chapter allows a larger theoretical framing and basis for the research-creation. These readings came before and during the creation process and framed the applied research investigation.

The research creation chapter expands on the questions posed in the introduction, and frames insights from the research creation process. It breaks down the process, considerations, and insights from this research. It is suggested to watch the "*Speaking Volumes"* (51:49 – 1:07:10) video before reading this chapter. This symposium articulated this chapter visually as a work in progress as the talk and this chapter writing was done in tandem. The conclusion summarizes key findings from the dissertation.

The remaining videos in the appendix were all created during the dissertation research creation process and frame the research in different ways. The artist talk *"STYLY x InterAccess XR Talk*" (36:50-1:01:40) is the first time I spoke of volumetric video research work in relation to my previous XR practice and was an opportunity to think through questions posed by other artists at the talk. It was through speaking about the work and thinking through concepts, that the importance of volumetric video as a means of bringing representation to VR over using prefabricated models was derived and added to the writing. The last video created was "*Research Creation Process"* and walks the viewer through the body of work created during the PhD program at York University. It situates how this research-creation component builds on previous artistic

research. Both videos provide a larger context within an artistic practice and act as a reference for the reader.

1. INTRODUCTION

This paper will document the methodology and approach of the applied research process and how insights were formed in the creation process. It will examine the role of volumetric video in immersive virtual reality environments. Like photogrammetry capture, volumetric video can be defined as point cloud-based depth recording of a moving image. Two central questions originally guided the applied research work:

Primary Research question: How can volumetric video inform the future of virtual reality and extended media environments?

Secondary: How can VR evolve from individual to social spaces?

Based on the process of applied research experiments, I believe volumetric video informs the future of virtual reality in three ways: in shaping what VR is beyond game mechanics into its own artistic medium; in how we tell stories as a mixed reality environment and in how we embody spectatorship as authors in developing immersive haptic media spaces.

To explore these questions, a series of experiments and explorations using volumetric video were conducted, primarily using Depthkit software. The dissertation includes both process documentation of the research creation process and the practical and theoretical evaluation of experiments. While the secondary research question guided my initial approach – and I still believe volumetric video might play a key role in the evolution of social VR – in the end it was not addressed in any depth due to the way my research evolved as a consequence of COVID-19 pandemic conditions.

As a way of introduction, I will start with the last part created to frame insights from the research creation process. The following is the voiceover script written for the volumetric VR experience, which was inspired by Jaron Lanier's various definitions of VR (Lanier 3-299) but emerged only after my own extensive engagement with

volumetric video. They are questions for artists to consider when developing a volumetric video VR experience.

How do volumetric tools shape the stories we tell?
How do the depictions of bodies in 3-D shape the narrative?
How many authors are needed to tell a story in volumetric VR?
What is the role of the camera in telling a volumetric story?
How does lighting create perspective?
What is the importance of scale in telling a volumetric story?
How do we develop a landscape for 3-D bodies in VR?
What is the function of a landscape in VR?
What is the function of architecture in VR?
What is the cultural significance of volumetric and other objects in virtual reality?
How do we show history in VR?
How do virtual material layers affect the story in VR?
Is the depiction of reality important in VR? How much reality is enough? Does volumetric video signal the real?
How does mixed reality affect experience with volumetric objects?
What is the importance of the glitch in volumetric? How does the glitch shape the experience?
How do we think about audience interaction when developing a story in volumetric VR?
How do you explore audience control in VR?
What is the role of volumetric characters in virtual reality? Are they avatars or actors?
What changes when we multiply the same characters in the experience?
How do we consider gravity in virtual reality? Does everything need to be anchored?
How many elements are needed to tell a story in VR?
How do we think about story genres in VR? Are these productively complicated by volumetric?
How do we direct the character's gaze in VR?
How do we use loops to shape the experience?
How do we think about timing and pacing?
How do author errors shape the experience?
How does technical error and glitches shape the experience?
How do we know when we've reached a story arc in VR?
How do we know when the experience should end?

The written component of the dissertation explores applied considerations in developing volumetric video for VR. Due to filming and developing during a pandemic, all research was done in a home studio and filmed and programmed by the author.

2. CONTEXT

This multi-modal dissertation focuses on research creation of haptic cinema and new ways of designing immersive story experiences with volumetric video. Recent technical innovation of volumetric video allows new definitions of haptic cinema to include self-gaze as explored by artists.

Sobchack's phenomenology of embodied spectatorship in cinema approach stresses the interactive quality of film viewing (Marks 149). Marks states: "If one understands film viewing as an exchange between two bodies -that of the viewer and that of the film – then the characterization of the film viewer as passive, vicarious, or projective must be replaced with a model of a viewer who participates in the production of the cinematic experience" (Marks 149). Marks elaborates on embodied spectatorship; she refers to Lacan's theory of the mirror phase but emphasizes that embodied experience is an important aspect of embodied spectatorship (Marks 151). She continues by situating her critique by saying, "Let me repeat that I am exploring sense of experience in cinema not to seek a primordial state of sensor innocence, but to find the cultural within the body. A valid critique of phenomenology is that it mistakenly believes that all of experience is accessible to consciousness; it just a matter of being in the present and perceiving" (Marks 152). She distinguishes the difference by presenting Merleau-Ponty's argument of embodied perception, of that of an examined perceptive view such as "the solidity of the ashtray, the roundness of the bowl, the blue of the rug – suggest that it is indeed easy for one to participate bodily in the world from the comfort of one's study " (Marks 152). She argues that the experience of the body informed by culture and that Merleau-Ponty influenced feminism, Foucault, and cultural anthropology. She indicates that "These studies show that our bodies encode history, which in turn informs how we perceive the world". (Marks 152) Marks examines the haptic relationship of film. She defines haptic perception as "usually defined by the psychologists as the combination of tactile, kinesthetic and proprioceptive functions, the way we experience touch both on the surface of and inside our bodies". (Marks 162). It is this definition that is expanded on through applied research. It will also expand to not only include the viewer but

consider the role of the artist as part of active embodied spectatorship of critical self-gaze as part of the creation process.

Haptic work could also create a feeling of distance, through miniaturization that evades the spectator's view, and instead pulls the viewer close (Marks 163). She argues that: "While optical perception privileges the representational power of the image, haptic perception privileges the material presence of the image" (Marks 163). She distinguishes the differences of both haptic and optic vision, but both are needed depending on the context. She states frankly, "it is hard to look closely at a lover's skin with optical vision; it is hard to drive a car with haptic vision" (Marks 163). She argues that haptic cinema does not invite an identification but a sensory-motor reaction that encourages a bodily relationship between the viewer and the image (Marks 164).

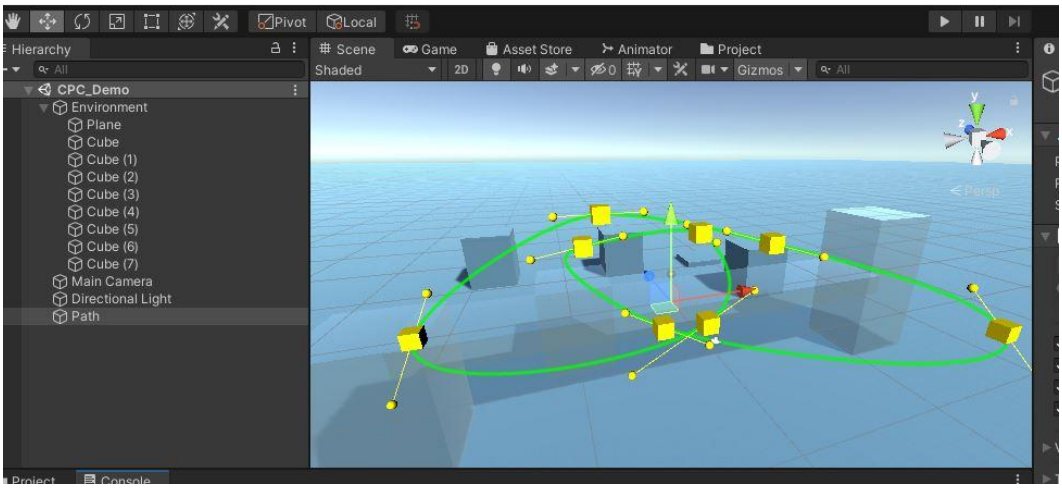

Image 1

In the applied research process, I explored different interaction systems to guide the interactive experience. I primarily used a tracked camera to move through the landscape allowing users to view the figures in a timed way. The viewer can see each figure up close, and at a distance. The role of author is clear by how the guided experience can pull the view close or far from the figures. The viewer can see the material presence of the image, and the breakdown or glitch of that image which allows for a deepened sense of haptic vision because the viewer will construct the gaps of the

body of the form to make sense of the image to make it complete. Even though a shell of a human is present, we understand that the figure was filmed live, and did not exist as a shell of a human. It is in the representation; we can see Mark's definition of haptic cinema evolve to include the participation in the production of the cinematic experience by reconstructing images based on embodied cultural knowledge.

Image 2

Marks notes that the term "haptic cinema" has a brief history, first described by Noel Burch as "Stylized, flat rendition of deep space" in early and experimental cinema (Marks 171). Marks also notes that Antonia Lant has similarly used the term "haptical cinema" to describe early films that exploit the contrast between visual flatness and depth. She refers to Deleuze's use of the term to describe Robert Bresson's *Pickpocket* where the sense of touch is isolated from its narrative function to create a cinema space. He writes, "The hand doubles its prehensile function (as object) by a connective function (of space): but, from the at moment, it is the whole why which doubles its optical function by a specifically "grabbing" (haptique) on, if we follow Riel's formula for a touch which is specific to the gaze" (Marks 171). Marks' position is to consider virtual reality as an embodied haptic cinema experience, a position that began to guide my own explorations.

To Marks, Deleuze's focus on filmic images of hands are unnecessary. She states: "Looking at hands would seem to evoke the sense of touch through identification, either with the person who hands they are or with the hands themselves. The haptic bypasses such identification and distance from the image it requires" (Marks 171). Marks indicates that intercultural film and video take advantage of haptic qualities of their media. Like the Deleuzian approach to time, "image strategies to prevent an easy connection to narrative, instead encouraging the viewer to engage with the image through memory. As fetishes protect their memories, haptic images can protect the viewer from the image, or the image from the viewer" (Marks 177). Marks states that haptics implies identification and that it requires a familiarity with the world the viewers know through more senses than vision alone (Marks 187). She states that her interest is not to anchor the definition into a psychoanalytic position but to suggest how haptic visuality works with the entire body: the "engagement of haptical visuality occurs not simply in psychic registers but in the sensorium" (Marks 188).

Maria Engberg and Jay Bolter challenge this concept of touch to consider virtual reality a disembodied experience. They ask the question "does cinematics contribute further to the decay of aura that Benjamin ascribed to screened films?" The authors invoke Benjamin's definition of aura, which gives a sense of distance no matter how near: "in tactile and proprioceptive terms – a work of auratic art never be touched by the viewer; it is always out of reach" (Bolter, Engberg 172). It is in this capacity that Enberg and Bolter take the position that virtual reality is a disembodied experience because things are in fact out of reach. I would argue what Engberg and Bolter do not describe is the close cultural identification that can occur through sight and touch in virtual reality, and I would argue that self-identification, perhaps particularly in the context of volumetric experiences, would be considered an embodied interaction.

Through this applied research process, it can be argued that Mark's definition of haptic perception can also be expanded, to not only include the embodied viewership of the viewer and image, but also the examination of relationship of the author's self-gaze in the creation of an image of self for display.

In her chapter "Drive(s)," Natalie Loveless describes the creator's and viewer's gaze as part of the research-creation process. She refers to Lacan's Paris lectures from 1964 on what he called "the gaze as object petit a," and Loveless describes the drawn object of the "petite a"' as a "researcher-creational thing-to-think" with (Loveless 80). It is the relationship between the "object petit a" the Lacanian "object-cause of desire" as the pedagogical lure that Loveless explores. She presents Lacan's second triangular schema, where the gaze, image screen, and subject of representation are intertwined and overlap. This overlap, the "realm of the imaginary (capture) interwoven with the "screen" of the symbolic" (Loveless, 92), is the lure. She discusses the "object a" in the filled with the visible gaze. It is this connection that Loveless describes as an enigmatic desire alluring point of light. Lacan ends his lectures with the question "to see, to want but to want and see what?" (Loveless 93). My volumetric capture investigation was initially inspired by the lack of representation of volumetric female images in the Unity and Unreal Engine storefronts. These storefronts have a "3D humanoid" section where digital assets can be purchase and brought into virtual environments. These assets are primarily character fantasy-based and do not consider cultural representation of real-world (particularly female) bodies experiences. Due to the COVID-19 pandemic and working in isolation, my research shifted, focusing increasingly on self-documentation and self-representation. Loveless' insights accompanied me through the research creation process of developing and theorizing recordings of myself, editing through self-gaze and ultimately building a virtual embodied immersive experience for myself, seeing myself multiplied.

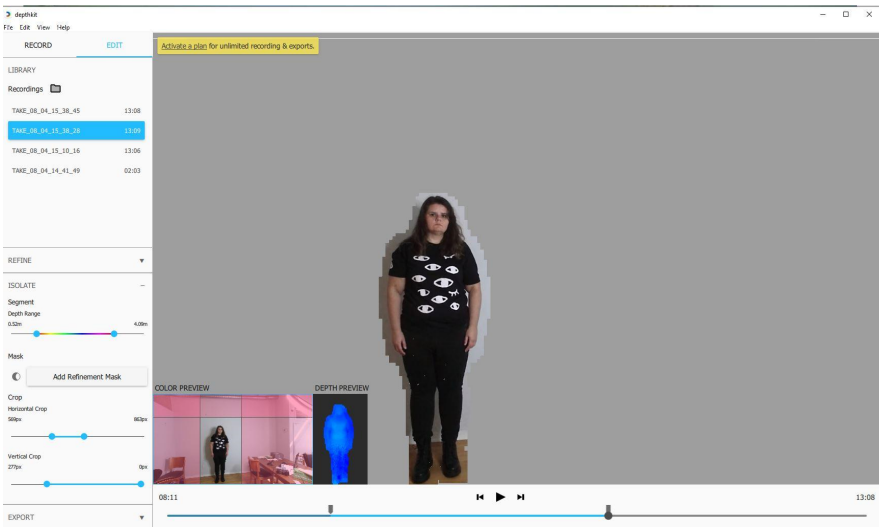

Image 3

When filming, the gesture of the body must be considered, and this would be represented virtually. Are they walking to a place? Speaking to the viewer? Acknowledging the camera and viewer? In my filming, I focused on the gesture of looking from left to right, and to left again to close a loop. This was translated later into Unity, with multiple figures looking at each other, with no gaze to the camera or acknowledgement of view viewed by the viewer or author. These experiments were directly informed by Loveless' writing on the creator's gaze.

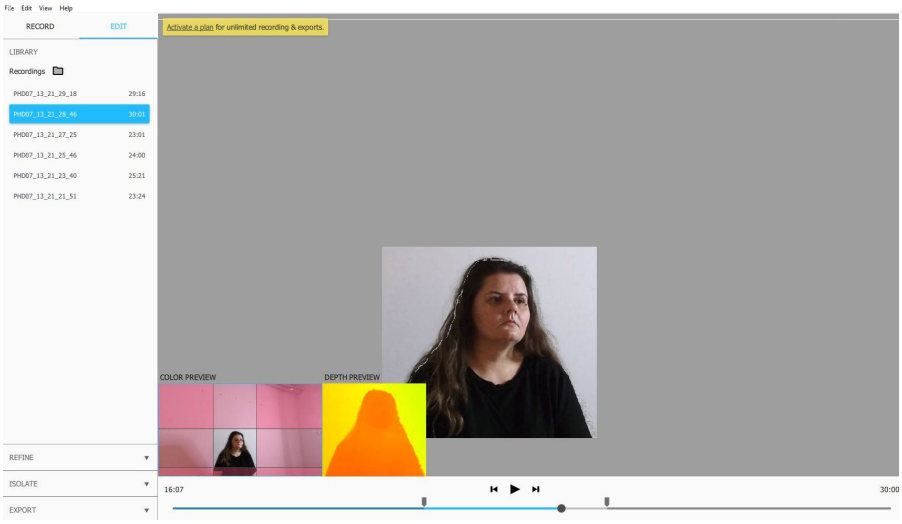

Image 4

By using volumetric video, we have new ways to engage of the viewer based on how the filmed figures look back and acknowledge being seen by the viewer.

In my research work, I wanted to build on this question and investigate the viewer's omnipresent gaze. However, working in isolation due to COVID-19, what I discovered was that the gaze is not in engagement with the viewer, but is a self-critical gaze in the creation process. I refer to this as the omnipresent self-gaze which is part of the applied research process. Working in isolation during COVID-19, I filmed, edited, and coded around images of myself for my own gaze in VR. This self-reflective study could be akin to my early arts training, drawing a self-portrait with a mirror for observation to build skill in observation drawing. Here, I gaze and replicate an image of myself to a translated digital artifact like VR. I would refer to this "lure" as the omnipresent self-gaze as one purposefully edited by the author for the author and viewer. It is this lure, that creates the self-edited process in order to "improve" the image for display. We learned, like that of "selfie" culture of social media sites such as Instagram, that in this creative volumetric video process, one's body could be documented, edited, and displayed as a document as a final three-dimensional digital artifact of engagement.

And, as with selfie culture, image, body, gaze, display and circulation of volumetric images is understood as having different consequences and affordances for different bodies. There is no universal subject here. While it is outside the scope of this dissertation to theorize gendered and raced bodied in volumetric capture, for example, it is not inconsequential that my self-representation challenges dominant volumetric representations of women that can be purchased online and that the only relations of looking that I've had to unpack involve only a single subject – myself - given the isolated conditions of creation and limited circulation to date. Thinking intersectionally about embodiment in the context of volumetric video is a promising direction for future work or to be taken up by others.

Here it's important to revisit the initial research questions: *How can volumetric video inform the future of virtual reality and extended media environments?* and *How can VR*

evolve from individual to social spaces? As I experimented, it became increasingly apparent how important it was to consider that within social interactions on screen, there is a feedback loop of self-gaze. We can see ourselves talking to another which is different than a real-world interaction of engagement.

Using the tools of a gaming engine, we can understand that applications can clearly be made for VR and AR. Using volumetric video can not only push narrative to show real life depictions, but with the glitch of volumetric video it can design new worlds. Henry Jenkins clarifies in his essay, “Game Design as Narrative Architecture", that not all games tell stories and may be closer to music or dance than to cinema (Jenkins 121). He gives the example of graphic games such as Tetris to illustrate his point. He discusses that game designers don’t tell stories, but design worlds and sculpt spaces (Jenkins 121). I would argue that volumetric video, even if used within a gaming engine to create a VR work, can exist purely as a designed world of sculpted space, and reflect real world aspects rather than rendered versions of real life that are algorithm based. These applications can potentially also make artworks more accessible for distribution beyond the gallery environment.

And yet, when thinking about volumetric video, we can see the promise in virtual reality and extended media environments as it applies to storytelling. Papagiannis distinguishes between AR and VR. She describes VR as a “special headset that blocks out your view of the physical world, trading the real world for a completely computer-generated environment” while AR allows users to be “more deeply immersed in, and connected to, the real world - the world in which we actually spend the majority of our time and attention” (Papagiannis 2). It can be argued that volumetric video bridges this gap. Volumetric video in VR incorporates aspects of the real physical world within a virtual closed environment. As seen in Image 2 above, the difference is that we understand that time is suspended in a virtual world, and rather than the volumetric video images appearing “live,” they appear more like an artifact or ghost.

Papagiannis outlines six storytelling conventions when it comes to AR storytelling (Papagiannis 78). The first, “virtual try-on,” makes the user be part of the story through prop, clothing, or augmented face filter experience. The second, “Hole in the wall, floor or table,” uses visual illusion to enter a new space. The third, “ghosts,” is unseen by the human eye and might even have events such as floating objects or illusions disappearing and reappearing within the augmented space. This is often associated with “glitch” based work in early iterations of AR experiences. This differs with the fourth storytelling convention, “living pictures,” which presents magic reality in AR, such as a “Harry Potter” AR experience (Papagiannis 80). The fifth, “x-ray vision,” is referring to a superhero power where you can see through objects. Papagiannis indicates this is often used in museums when looking at artifacts. The sixth, “3-D drawing,” is about users authoring their own reality in augmented spaces. This is often used with hand-based gestures such as colouring and drawing in virtual spaces. Through Papagiannis’ definition, it can be argued that volumetric video in VR situates itself between the third and fourth conventions. It could be argued that the “glitch” appears as a ghost, but it can also be argued that it removes a sense of real time because we can see the fabric of the technology. It is in seeing the pixeled images that we understand we have a fragmented view of a whole image. It is in this fragmented image we understand that information is lost in the representation. It also can appear as a “magic reality” of living pictures where we are combining real life images within artificial environments. Therefore, volumetric video VR would be like a subset of AR storytelling within a VR environment.

It also important to note how volumetric video and photogrammetry are also used to document reality. Steyerl describes how 3D scanners are often understood as a “technology of truth”, often used for police work and investigations for documentary evidence. In the scanning process, however, missing information and fractional space can be created when constructing 3D scans. Steyerl argues it is in the repairing of such images that fiction is inserted in the image to reconstruct these spaces.

Steyerl also describes how the representation of a 3D scan does not produce bodies or objects but folded surfaces. It is these surfaces that are wrapped to create a volumetric image. Steyerl suggest that depth is created by folding a surface onto itself by folding and stretching it. It is also noted that the surface carries cultural imprints that are social, political, and technological in nature. She argues, then, that surface is not only appearance but also a representation of the folding of historical and social information. It is in this way that it is important to consider the creation and editing of the image process when developing volumetric images and how this editing changes the story from fact to fiction.

When considering the frame in emerging media, one must consider the implications for the viewer within its immersive space surroundings. It is in this fundamental way that cinema and VR are different. Cinema is fixed in its storytelling and point of view, where ultimately VR allows the viewer to shift visual perspectives. Bolter refers to the writing of Anne Balsamo who has described virtual reality as a new staging of the filmic eye: “in most VR programs, a user experiences VR through a disembodied gaze, a floating moving 'perspective’ - that mimes the movement of a disembodied camera eye” (Bolter 248). This, Bolter states, “is a familiar aspect of what may be called filmic phenomenology where the camera simulates the movement of a perspective that rarely includes a self-referential visual inspection of the body as the vehicle of that perspective” (Bolter 248). There is a disembodied gaze that happens in VR environments. It is not only what we see, but how we see it. In VR environments we see that representation of objects can be infinitely reproduced. It is a world where trees can be virtually created and copied algorithmically into infinity. It is not only the content that can be reproduced but multiplicity is also the experience itself. Anne Friedberg discusses Bergson’s point of view of Eisenstein and the concept of “multiplicity”, that the viewer can be in two places at the same time. Ultimately indicating that “in ways that neither Bergson nor Eisenstein could have foreseen - to inhabit, in a virtual sense, two or more spaces at once, and equally, two or more times” (Friedberg 146). It is within this discourse of multi-place, where Friedberg refers to Deleuze's concept *of* “radical elsewhere” and that “doors, windows, box office windows, skylights, car windows,

mirrors are all frames" (Friedberg 241). Friedberg considers these as closed systems connected through these frames, and beyond the framed screen is a "radical elsewhere" and there are always things beyond and outside of the frame. It is in this discourse we can understand how telepresence and VR can be understood and evolved in new ways.

In their chapter "Mobile Cinematics", Bolter and Engberg argue that the most innovative part of the interface offered by VR is proprioceptive. They argue that "proprioception contributes strongly to the sense of immersion and presence" (Bolter, Engberg 168). They also discuss the figural and the panoramic in relation to the perception of scale, arguing that scale has been an important variable for cinema for the large screen and television for the small screen but with immersive mobile VR it instead creates an intimate space. They argue that AR and VR applications are "correlated to specific positions, movements and actions of the human body" which are then translated to spatial scale and perception (Bolter, Engberg 170). They build on the definitions of Montello, Barba and MacIntyre noting that *"*the scale closest to the user is labeled as the "figural scale*"* (Bolter, Engberg 170). It is this taxonomy of scale that users can experience objects through manipulation, see vistas beyond their reach, and turn their own bodies for 360-degree viewing. It is this taxonomy of scale they believe is unique to VR and different to that of cinema or television.

Although they give the example of mountains and skyline out of reach, my research-creation experiments lead me to argue that it is less the aura of the unattainable reach but, rather, the coexistence and sense of presence one needs to consider when building volumetric video in VR and playing with scale.

In my work, the viewer is embodied in a world with multiple images of the author. It is through the multiplicity that we understand that it is not a depiction of a real world but a depiction that uses elements to ground us into a reality. Patrick Lichty indicates that VR may represent an "intermediary position between the objective and the non - by incorporating embodied, immersive experience" and that in this way, it is an inverse of

relationship to that of performance art because the “interactor performs the artist’s installation space (Lichty 49). Similar to Lichty’s definition, we understand we’re in an artist’s space that is moving us through a landscape in a strategic way. It is in this way that volumetric video brings us to think of VR as performance, installation, cinema, and gaming environments in unison as an embodied experience by the viewer. In using volumetric video, we can see familiar cinematic representations of characters, but what makes it fundamentally different is the way we interact with the screen and the medium itself. As an artist or maker, one could infinitely multiple the cinematic figure as we would an object or parts of the landscape. It is in this way we can see the separation of representation of the image within the digital context as the filmic image can be replicated and manipulated in an infinite way. This can fundamentally shift how we understand VR through the use of volumetric video.

How can VR evolve from individual to social spaces?

This applied research explored how a VR environment creates a constructed reality through synthetic landscapes and volumetric video. Although industry tools were being used, the spectrum of research focused on artistic research creation. It is not created for the purposes of mass distribution but using commercial tools for artistic development at an experimental prototype stage. Adorno and Horkeimer reflect on mass consumer models of content creation and indicate that film has taught spectators what to expect and how to react automatically (Adorno, Horkeimer 4). As the volumetric video telexistence work was in an early prototype stage, it is premature to come to conclusions and it is speculative to conclude how VR can evolve from individual to social spaces. Although the author can visit the online space, it is unclear how duration and how multiple viewers experiencing an online volumetric space changes from a single immersive VR experience.

In his writing on virtual reality, Lanier does argue that the body is adaptive, and that a new reality is constructed despite the improbability of the content in which one is immersed: “we are not fixed targets, but creative processes” (Lanier 50). The invention of volumetric video allows users to see something familiar from their real-world

experience in VR. It gives audiences familiar (if sometimes uncanny) ways of reading images and filmed characters in a new environment and in the context of a new disembodied way of seeing. It is in this way that volumetric video in VR is fundamentally different, as previously it was only the digital rendering of how the computer envisioned a person could look like through shapes and animated form, rather than a digital translation of a person through a filmic apparatus. Bettina Back refers to the viewer's body as a new frame. (Back 21) She indicates that a prerequisite for media of all VR worlds is "their genuine performativity" (Back 21) and that the viewer's place in these worlds shifts into their physical memory. That is a sensual-haptic body perception. She is interested in the sensorial moments within interactive real-time simulations.

As stated in the introduction, Marks also refers to the writing of Vivian Sobchack who argues that cinema is not an illusion, but an extension of the viewer's embodied experience. (Marks 149). Marks elaborates on embodied spectatorship; she refers to Lacan's theory of the mirror phase however emphasizes that embodied experience is an important aspect of embodied spectatorship (Marks 151). She continues by situating her critique by saying *"*Let me repeat that I am exploring a sense of experience in cinema not to seek a primordial state of sensory innocence, but to find culture within the body. A valid critique of phenomenology is that it mistakenly believes that all of experience is accessible to consciousness; it just a matter of benign in the present and perceiving*"* (Marks 152). Marks distinguishes the difference by presenting Merleau-Ponty's argument of embodied perception, of that of an examined perceptive view such as *"*the solidity of the ashtray, the roundness of the bowl, the blue of the rug – suggest that it is indeed easy for one to participate bodily in the world from the comfort of one's study*"* (Marks 152). She argues that the experience of the body informed by culture is implicit in Merleau-Ponty indicating that that *"*These studies show that our bodies encode history, which in turn informs how we perceive the world*"* (Marks 152).

I carried this idea into my research-creation practice. Although VR can embody a point of view and a new perspective, we cannot neglect the embodied spectatorship and unique point of view a viewer can bring, especially in telexistence interactive environments, even if they are in an avatar skin. Marks states that haptics implies

identification, that it requires a familiarity with the world the viewers know through more senses than vision alone (Marks 187). She further states that her interest is not to anchor the definition into a psychoanalytic position but, rather, to suggest how haptic visuality works with the entire body: the "*engagement of haptical visuality occurs not simply in psychic registers but in the sensorium."* (Marks 188). This builds upon earlier definitions that allow us to think about embodied engagement, especially deciphering glitched volumetric images. Telexistence is discussed in Roquet's summary of Susumu Tachi's research. Tachi was a student of Takemochi Ishii, pioneering Japanese research in the field of human-computer interaction in the 1960s. Tachi concluded that while *"*telepresence has tended to focus on teleoperated robots in real-world locations, often positioned in contrast to VR's focus on computer-generated worlds, telexistence can target either real or virtual spaces (or a mixture of the two) essentially combining two fields*"* (Roquet 40) Tachi's focus was on how to incorporate stereoscopic head-mounted displaces as telepresence experiences.

Roquet brings an additional perspective to virtual reality development, and it is what he refers to as the "Telepresence enclosure" which is the mediated privatization of presence itself (Roquet 33). Roquet articulates Aizu Izumi's speculation that in the twenty-first century people were allowed to physically move freely; however, the twenty-first century *"*Might come to an era where people have their will to move freely while they are physically in one place" (Roquet 34). Roquet further notes that VR demands physical enclosure in exchange for perceptual mobility. Here Roquet considers telepresence's shift to telexistence. Telepresence, he argues, promises to extend human perception across inhuman distances, unlike traditional VR approaches that locate the space of interaction entirely within a computer (Roquet 36). He speaks about the human gesture and how it can effectively take place in two locations at once: one, initiated at the site, the other with the gesture effects registered in another place entirely. He further defines telepresence from describing what was being transported here, *"*physically embodied agency at a distance (the original Latin praesentia, appropriately means 'being at hand.*"*" (Roquet 36) This is critical in the mixture of real and virtual experience because it allows not only the disembodied gaze but connects an artificial place with a real physical place. It is also critical in understanding the evolution

of telepresence and its shift to telexistence which is a mixture of real and virtual experience. It is speculative how volumetric video AR might evolve as a telexistence mobile experience because place and distance may shift with time and geolocation. It is unclear how this change of physical environment beyond a screen may contribute to the interaction.

We see that these ideas of telepresence and telexistence have evolved not only in engineering practice but practice-based research. In her 2020 collaboration with Eyebeam, Valencia James poses key questions to formulate her artistic research work: *How can machine learning inform dance and movement? How can we create a multisensory inclusive experience in XR? How can creation in immersive web space be an act of liberation for marginalized voices? What is the decolonized future of the Internet? (*James 2021*)* James was interested in the radical intervention of digital spaces co-creating the volumetric toolbox where she created live events using the toolbox and Mozilla spoke hubs. I believe VR chat and WebXR spaces are promising in this regard. It is an opportunity to modify existing ways of communication and creation to be more inclusive. James created internet spaces where participants appeared as avatars that can walk around the volumetric video performance space. Unlike Tachi's development of commercial systems, James and Eyebeam have created an open toolbox; accessibly priced for artists and designers to experiment with and evolve the platform in an open lab environment. Working with existing artist-made tool sets was a starting point of creative research for my own research work, including that developed by James and Eyebeam in 2020. Although the telexistence aspect of the research was not fully developed, as tools become more accessible this is one area, I would like to further explore them, especially with the inclusion of live spatialized sound that is unique to an immersive VR experience. I would also like to expand on the list of questions for consideration when developing volumetric video for VR that was developed in the research work to consider techno-social spaces.

3. RESEARCH CREATION

Loveless indicates that the term research-creation works in tandem with other terms such as practice-based research, creative-praxis, arts-driven inquiry, arts-based research, and artistic research (Loveless 4). She refers to Frayling's model as an example of the research-creation process.

Frayling's methodology is based on the Royal College of Art model of Art and Design Research as published in the inaugural edition of the Royal College of Art Research papers. In Frayling's writings on research in art and design, he indicates there are three major differentiations on art and design research: *Research into Art and design*, *Research through art and design*, and *research for art and design* (Frayling 5).

The first, *research into art and design,* focuses on historical research, aesthetic research, and research into theoretical perspectives. The second, *research through art and design,* focuses on material research, development of work, and action-based work as a way of a practical experiment in the studio. The third, research for art and design he argues is complex and conceptual. It is "Research where the end product is an artifact-where the thinking is, so to speak, embodied in the artifact, where the goal is not primarily communicable knowledge in the sense of verbal communication" (Frayling 5). This research approach followed the "*Research through art and design"* process. It looked at how research creation led to new insights into the field of volumetric video in VR environments. The following is a walkthrough of artistic considerations that led to the questions articulated in the voiceover outlined in the introduction.

EARLY STAGE APPLIED RESEARCH

In October 2018, I participated in a volumetric video incubator at Dames Making Games in partnership with Depthkit. Depthkit is a software that translates depth camera images that can then be used in VR environments. A multi-camera set up can provide a 360-degree capture of a person or object. Preliminary research was done using Unreal

engine, a beta testing environment, in collaboration with the makers of Depthkit. Over the forty-eight-hour workshop, short sequences were filmed and edited into short loops and placed in a synthetic landscape. Depthkit no longer supported Unreal Engine beyond the beta testing phase. I continued research with Unity and Depthkit plugins during the research creation process of this dissertation.

As seen in Image 5, an image of myself was filmed using an Xbox Kinect camera. We see the filmed image and the depth image translated below. This was initially imported into a beta version of Unreal plugin supported by Depthkit which has since been discontinued. As with early beta software, often versions are discontinued, and works will remain at a prototype stage.

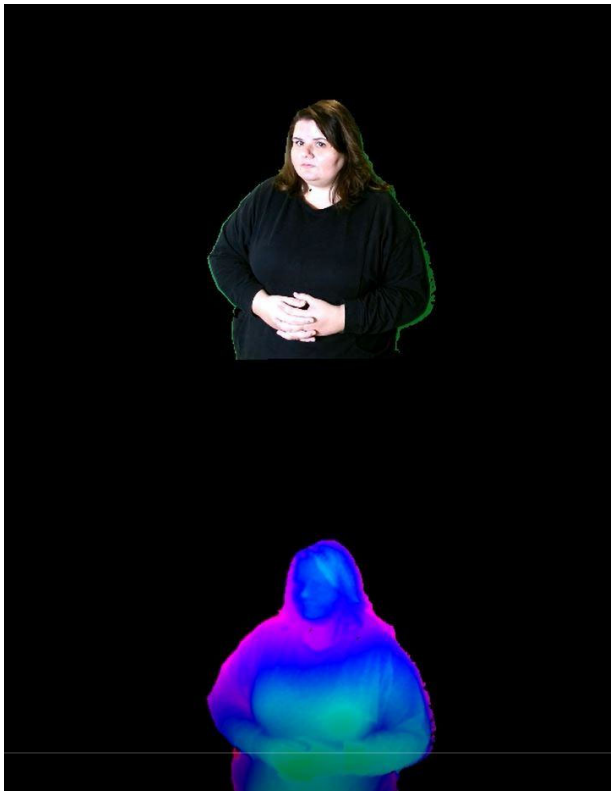

Image 5

Image 6

Part of the initial applied research focused on emerging tools in volumetric video and VR. Technical research was done using the Unreal Engine and Depthkit with volumetric video being recorded with the Xbox Kinect camera.

The drawbacks of working with the Unreal Engine included the limited number of videos that could be brought into the interface. The disparity of the video-rendered image was problematic compared to a highly rendered environment. The digital environment was created by generated materials texture on a wireframe textured landscape. Lighting affected the volumetric video. Although we see reflections of the sky, for example, on the landscape, we are unable to see a clear reflection. It was an important consideration going forward to think about how the initial video is lit to bring it into an electronic environment. These insights were key prior to the filming of the next phase of the applied research process.

LIVE SPATIAL SOUND

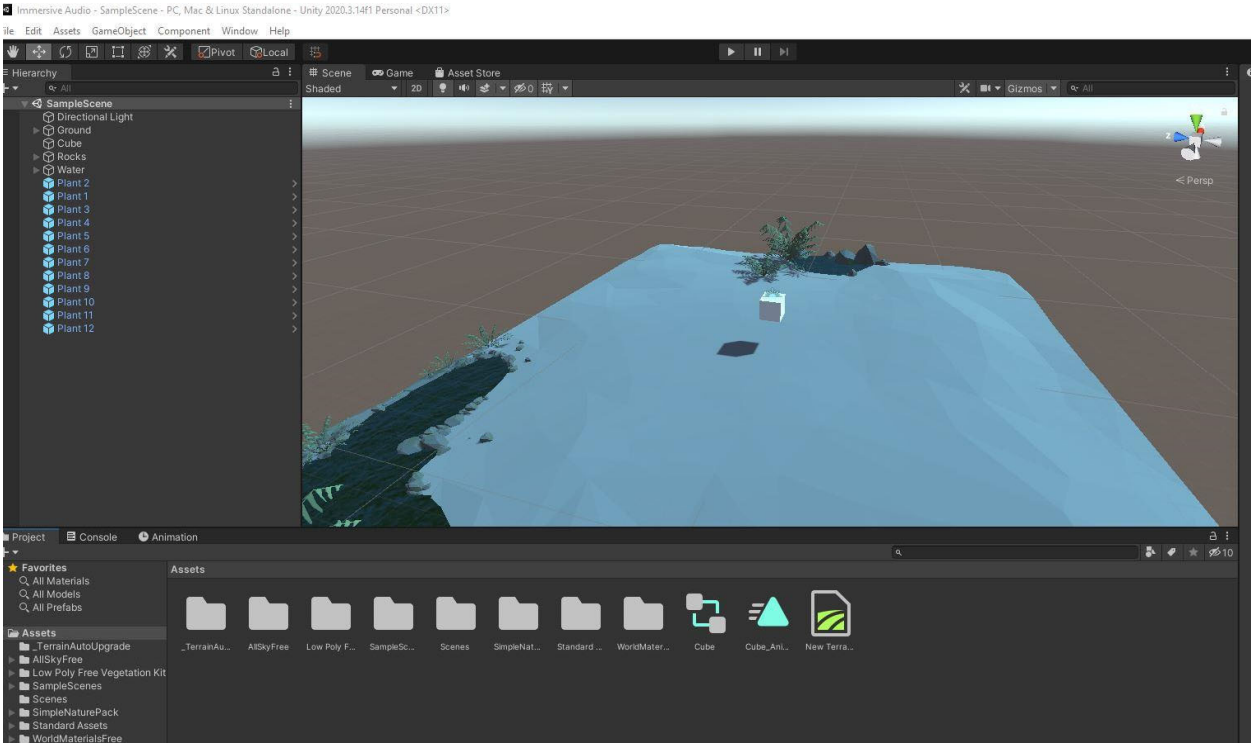

Image 7

In November 2020, I took an online workshop with audio artist Dierdre Morrison in immersive audio and experimental sound design using Unity. I did this to get oriented to the Unity platform which differed from Unreal Engine, and to consider new tools in designing sound for VR spaces. I was also introduced to Ableton, which can be used as a bridge into Unity for live music composition. Like silent cinema with accompanied piano, I considered using Ableton to enable me to design live sound to add to a volumetric video VR space. Early experiments were created, but ultimately a voice-over narrative was used in lieu of live digital orchestra as a method of storytelling.

The takeaway was that the audio composition could not be replicated, so no viewer would have the same experience. Live sound was discarded to focus on volumetric video and the voice of the author. Although the experiments had limited success, it was an important companion to the research rather than using digitally generated instruments. Also, the processing of both Ableton and Unity created a significant lag of

the computer processor which limited the number of volumetric videos and compromised the audio composition.

WEB MOTION CAPTURE

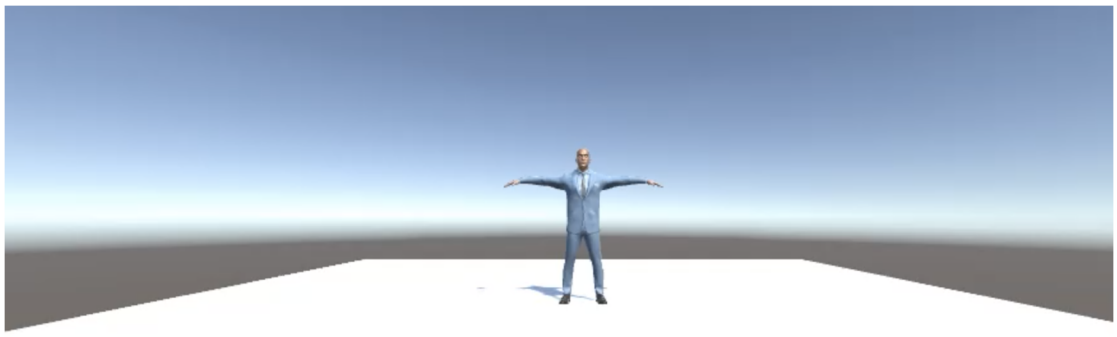

Image 8

In January 2021, I took an online workshop with artist Freya Olafson to learn how to use motion capture with a webcam for import into Unity using a web tool called Radical. Taking the workshop was both a way to advance my skills with Unity and was also an opportunity to consider how one could film a body to use as a skeleton to bring into VR. As seen, above, my body movement gesture was scanned via webcam, and then a prefabricated figure was mapped onto the skeleton structure. The workshop focused on how photography can be used for immersive VR environments. Although an interesting approach, volumetric video cannot be mapped onto different skeletons, which meant that this research direction could not inform my understanding of how volumetric video might inform the future of virtual reality and extended media environments so I did not to pursue this research route at this time. I did consider how gesture could inform the research as part of this process, and if it is unique to each body such a cycle or hand movement.

MOZILLA SPOKE HUBS

In February 2021 I attended a talk of Valencia James' research work with Eyebeam. James was interested in the radical intervention of digital spaces and co-created the volumetric toolbox where she created live events using the toolbox and Mozilla Spoke

Hubs. She created internet spaces where participants appeared as avatars and could walk around the volumetric video performance space. James and Eyebeam have created an open toolbox; accessibly priced for artists and designers to experiment and evolve the platform in an open lab environment. James invited me to test out the process and granted me permissions to work on her Mozilla spoke hubs server and use the volumetric toolbox. Working with her existing artist-made tool sets and others developed by James and Eyebeam in 2020 advanced my research by considering how volumetric video spaces could be social and outside of a VR headset.

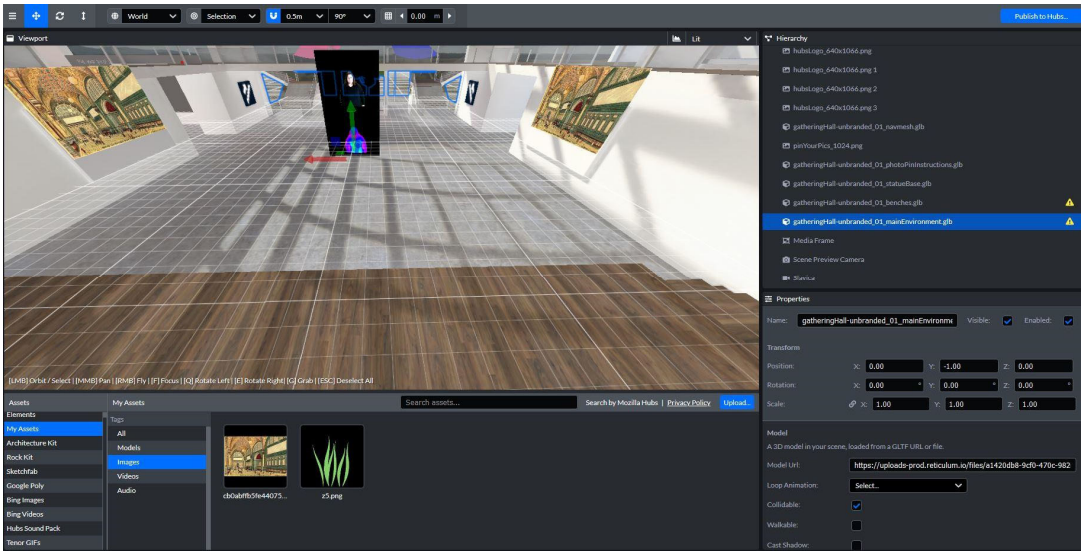

Image 9

As with any emerging tools, there were limitations. Unfortunately, no drivers are written for the Kinect camera, as James' approach uses a camera made with an Arduino kit. There is also no documentation on how to use the tools as they were in process of being created. I started to embed placeholder screens of volumetric video I filmed in a previous Depthkit workshop and started to consider developing virtual waiting rooms. I was interested in how we host spaces as artists and authors, and where people gather in physical space. I started looking at early twentieth century waiting rooms for women in train stations in New York. I was interested to see how social construct of spaces was created for women, and what this would look like today.

Ultimately this approach was abandoned due to technical limitations of a new toolset. It did inform the next phase of my project, and I considered if multiple figures are in a VR environment on a loop and if this could create as similar waiting or holding of space. I decided to explore this avenue further to deepen my understanding of volumetric video in VR.

TOOLS

In July 2021, I began filming myself with an Xbox Kinect depth camera in my studio space at home. Due to health regulations in Toronto, I filmed the work on my own with a single Xbox Kinect camera that was translated with the Depthkit software. The Depthkit software records a depth image, and allows it to be brought into Unity, a commercial game engine to develop a VR work. The camera, originally developed to be used for motion-based commercial games for Xbox, is appropriated for volumetric filmmaking.

Image 10

As shown in Image 10, the following equipment was used in the applied research phase:

- Gaming PC laptop
- Xbox Kinect 2 camera
- Tripod
- Kinect PC Adapter
- Three button mouse
- External hard drive
- Microphone
- Quest 2
- Quest PC Link cable
- Ring light

The software I used during the creative process:

- Depthkit
- Unity
- Ableton

The tools available are commercial tools used for development of video games. Unity is a gaming engine with various versions and iterations. The advantage of using an engine is that the mechanics for interaction are in place; however, with updates on the software, any works derived may be obsolete if plugins are not supported in updated versions.

In the writing of this paper, screen capture software was used for video walkthrough and photoshop was used to edit photographs to crop for documentation of screenshots.

Part of my research was understanding what happens when filmed physical representation is brought into virtual environments. I wanted to understand how to create spaces around virtual represented bodies.

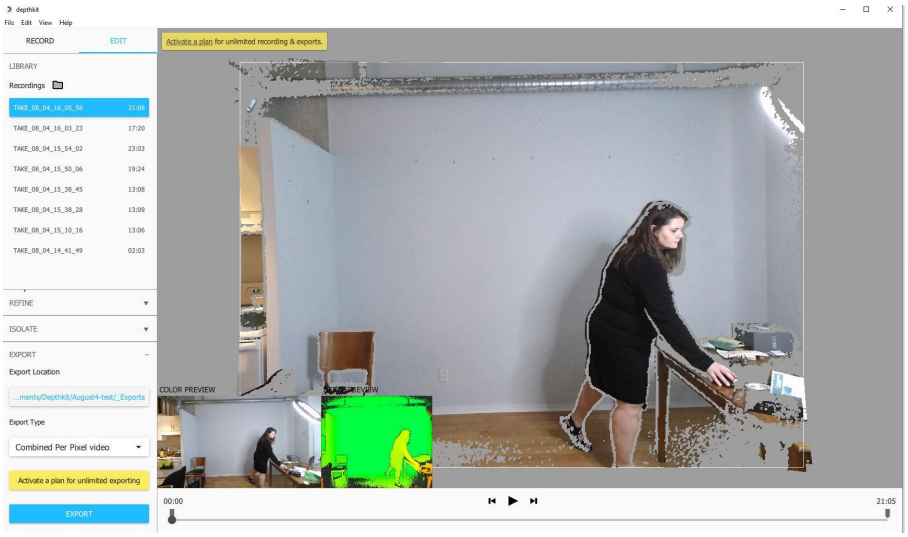

Image 11

How much physical space is needed to physically film oneself in real space intended for virtual space? In my previous work, I filmed above the waist with a six-foot green screen backdrop. Without those tools, I had to film far enough away where I could film my whole body to bring that footage into a virtual Unity environment. The area was completely cleared to film. As the Xbox Kinect camera cannot zoom in, or out, it was required to be fixed on a camera tripod.

When filming bodies for volumetric video there are several factors, one must consider:

1) Developing timed video loops
2) Depth and angle of image
3) Narrative based on body movement
4) Lighting

First, is the timing of the clip. In the creation process, a series of clips are made, and the image is broken down into three files. A text file, a video capture file, and a photographic depth image. I filmed a series of sequences that could be brought in as various timed loops. The loop will exist in the narrative and will give the reflection of a real-world experience in VR. If the clips are of different lengths the loops would change the pacing of the VR interaction. As stated earlier in this paper, the filming was informed by Natalie Loveless description of the creator's and viewer's gaze as part of the research-creation process. It was discovered during the research creation process, that the gaze

is not in engagement with the viewer, but a self-critical gaze in the creation process. I refer to this as the omnipresent self-gaze which is part of the applied research process.

Marie-Laure Ryan outlines nine various narrative structures for virtual reality. The model developed here most closely resembles her seventh model. In her seventh model, Ryan describes “The directed network, or Flow Chart”. In this model the reader may hit a series of dead ends and feels like they are moving in circles. Ultimately there is a horizontal progression that moves the user through the narrative with a possibility of different endings. In the eighth model, “The hidden story”, it allows users to discover a pre-history to a narrative or game world. She compares this story structure to “Myst” (Ryan 253). She states that “this model consists of two narrative levels: at the bottom, the fixed unilinear, temporally directed story of the events to be reconstituted; on top, the atemporal network of choices that reader-detective’s investigation” (Ryan 253-4).

The understanding is that the story is woven together by the user’s attempt to reconstitute the underlying story. This story is guided by a narration of how to consider the image, and how to consider the artistic creation of an image for volumetric video. In developing a consideration as articulated in the introduction, it creates an almost pre-story of a guided directed narrative through a guided rail-based interaction for the viewer. We are moved through the story both being told how to consider the image and are virtually moved through the story and shown what to look at.

When artists are developing interaction models and narrative with volumetric video and VR, it is also important to understand the nature of the spectacle and how the volumetric video part of a whole structure of VR is and cannot be completely separated. In Brenda Laurel’s essay “The Six Elements and the causal relations among them”, she envisions interaction with computers in theatrical terms. She references Aristotle’s ideas about drama and that a finished play is an “organic whole” (Laurel 564) where the whole organism is more than the sum of its parts. Aristotle identified six qualitative elements of drama and the relationship among them. In terms of formal hierarchy, there is action, character, thought, language, pattern, and enactment. (Laurel 564) For Aristotle, as

Laurel articulates it, the fundamental material element of drama is “spectacle”, or “performance”. In human-computer interaction Laurel argues that the whole “action” is collaboratively shaped by system and user, and the “spectacle” is the sensory dimension of the action being represented either by visual, auditory, kinesthetic, and tactile elements. (Laurel 565) Laurel refers to the early production of Lanterna Magica in Czechoslovakia and an “interactive movie” in the Czech Pavilion at the 1967 World Expo in Montreal. Laurel writes about the perception of patterns as sensory phenomena and a source of pleasure for humans. Aristotle described this as “melody”, and that the interpretation is that “spectacle” is the visual dimension of “melody”. Laurel discusses the role of beauty that drives Aristotle’s criteria and that the beauty of form and structure “can approach that of natural organisms in the way parts fit perfectly together”. (Laurel 570) She parallels this to human-computer activities as beautiful and criteria should be used in deciding what the person should do in a computer-based agent system as a course of action. What volumetric video and VR adds is depth to an image, and how in a multi-camera environment we can develop the melody of the organic whole to include haptic cinematic perception.

In the image creation process, depth must also be considered. In a multi- camera environment, images are taken from various angles and placed together. Given a one camera scenario, artists will need to review the depth of the image and what is exported as a clip. The image can be cropped, and then exported removing the background from the image. In removing one background, you will ultimately be replacing this for another as part of a new world view for VR.

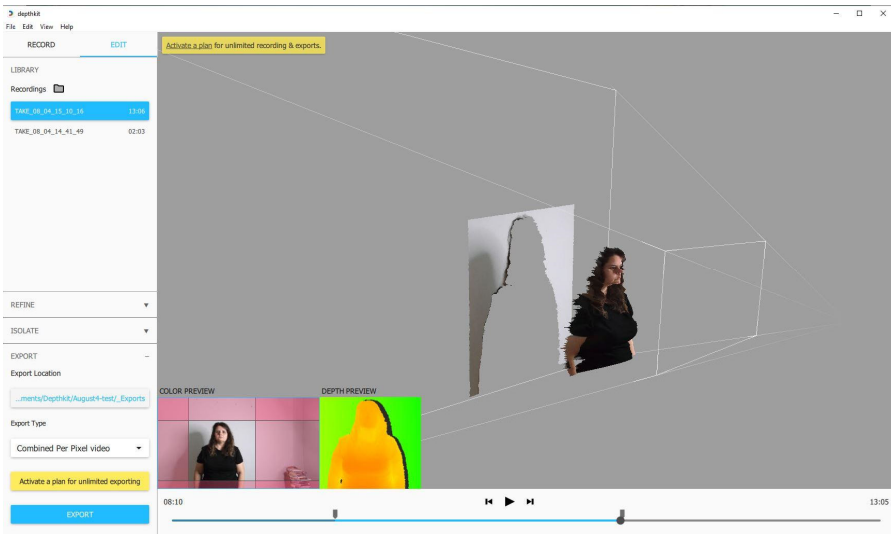

Image 12

I would argue that there is a multiplicity of images in a 3D scan. Even in a single camera depth image, the output of the recording will be the original image, and the depth image. This sequence was brought into Unity, and able to map as 3-D virtual objects with the Depthkit plugin within Unity. Materials cannot be added; however, the image could be scaled within the virtual environment.

Filming with a single camera gives you a 3-D image of a body, but the body translates into almost a shell: A fragmented, glitched presentation of a human. This allows us to think about how to construct a landscape around the figure, allowing the users to reconstruct the remaining image on their own.

Lighting is a specific requirement that must be considered when developing volumetric video for VR. A screen for visual feedback was needed to see what is filmed as the depth camera does not have a viewfinder for framing. The camera has a sensitivity to light that is needed to capture the depth of an image. Most sequences that were filmed successfully were filmed during the day and all at the same time for consistency. One ring light camera was used but it didn't evenly distribute light for a 3-D image. This cast shadows that limited how the image could be placed and lit in the virtual environment and affected how shadows were developed for other objects in VR.

These issues, although practical, allowed me to understand larger theoretical questions about interactions with volumetric video in VR. It was important to consider how assets brought together must work in a harmonious melodic way to make the interactive experience cohesive for the viewer. These technical refinements are part of developing a technical system to create user engagement through interaction.

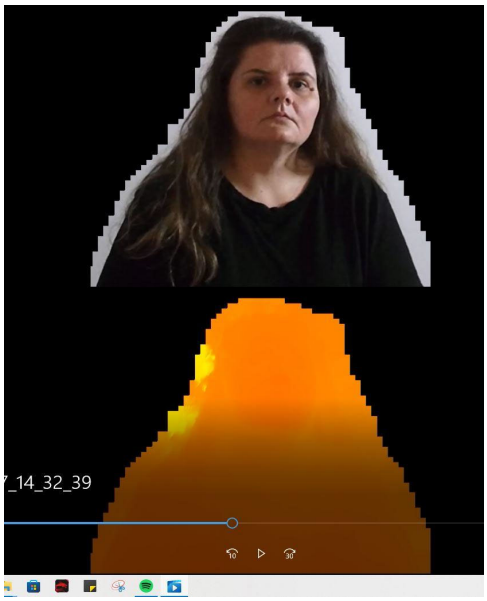

Image 13

DEVELOPING VIRTUAL REALITY SPACE

Once the sequences were edited, they were brought into Unity. This was the first time I saw the volumetric video sequences in a virtual environment. The first thing I needed to consider is how the videos performed in the virtual space. Did the loops glitch? How many videos can work successfully without lag? I chose to line up the figures into sequences where they were looped. This creates a “looking sequence” where the video profiles were looking at each other in a looped way.

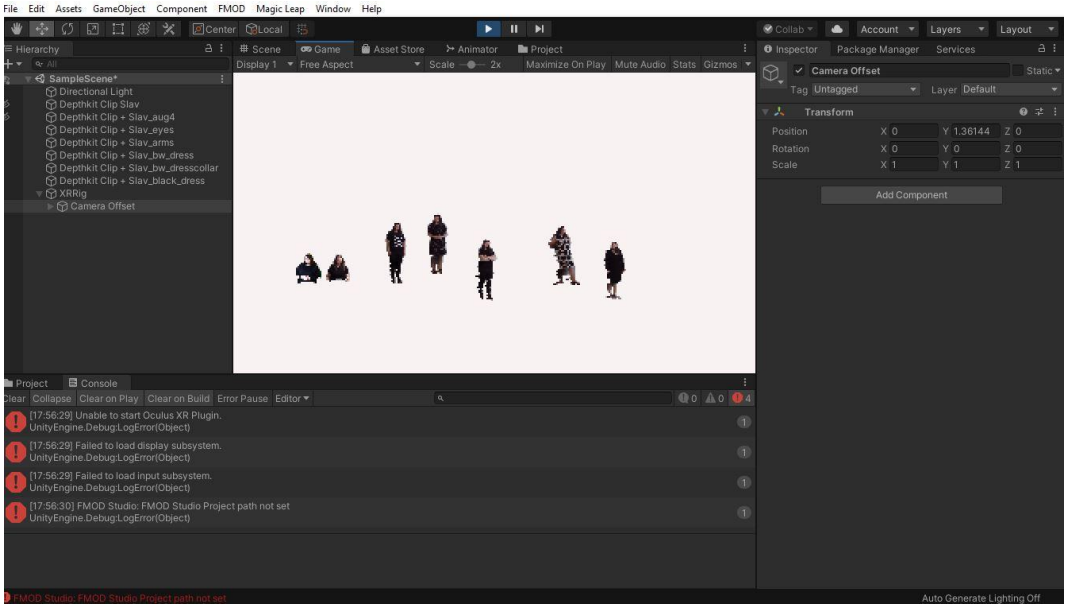

Image 14

When in “play” mode, we can see what the space looked like with a fixed camera. Although filmed in 3-D, with a static camera, the scene almost felt like a theater stage with characters waiting to perform. Although it sets a scene, it lacks the depth of a virtual reality world.

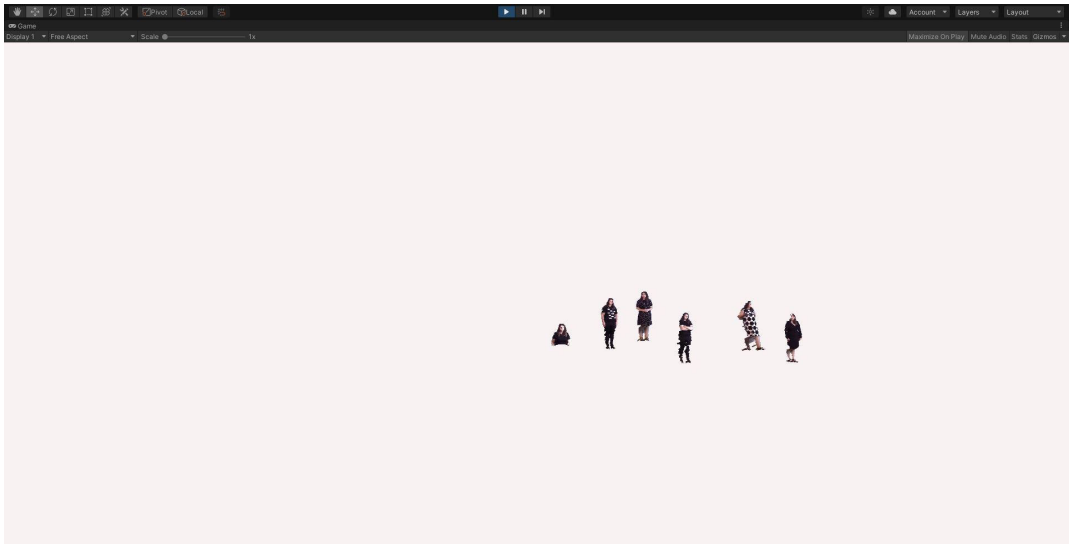

Image 15

One of the first things to consider is the camera placement in relation to the figure. It is also important to consider surrounding landscape, and the scale of the figures in relation to the landscape. This differs to traditional film because not only are we filming the characters, but we are also setting the filmic frame for the interaction with the viewer. These cameras work in unison. First capturing the gaze of the filmed character, and second camera directing the omni-present gaze of the viewer. It is in the research creation process that new questions begin to form, which ultimately became the artistic considerations for building volumetric video which was used as a voiceover as outlined in the introduction.

DEVELOPING LANDSCAPE AND ARCHITECTURE

When developing a virtual environment, various digital elements need to be considered. When I began the world design, I believed that the design decisions would be no different than architecture design in the real world, and one must consider how to navigate the world and what functionality is needed. However, I found, through experimentation, that digital spaces differ in that those architectures might not serve any function at all. A camera can move through walls of a virtual building without obstacle. The building could be also purely decorative in nature and serve no purpose as a shelter. The architecture itself could be designed in 3-D modeling software and brought into Unity, or architectures and landscape assets could be prefabricated and brought into the virtual environment through the Unity store. The consequences of prefabricated models are that the model itself cannot be changed; however, scale and material effects could easily change based on the artist’s decisions.

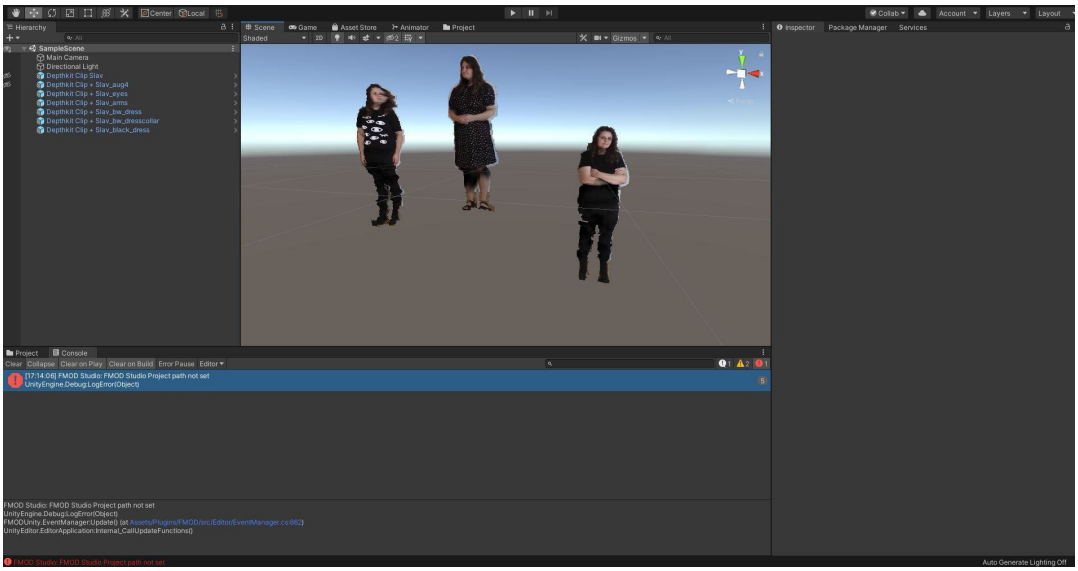

Image 16

The cultural significance of objects and architectures imported from other virtual homes and transplanted needs to be considered. Material textures can change. However, if the homes do not function as shelter, does an abandoned home shape, and effect the narrative as part of the built atmosphere?

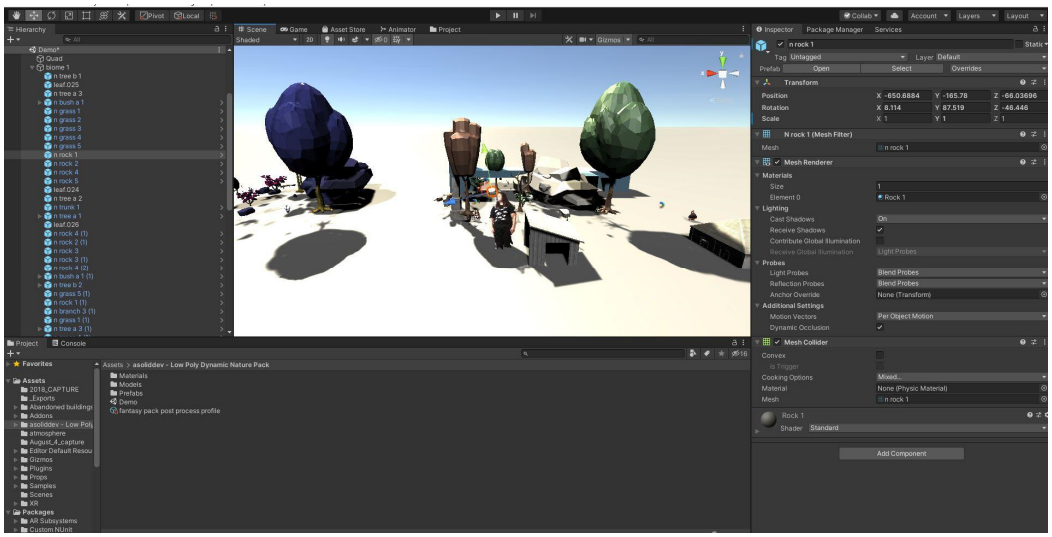

Image 17

It is also important to consider if ‘reality’ – here understood as invoking the physical world - is an important factor when building environments for virtual reality. Using

volumetric video brings a reality into the scene, but it can live in any environment and logic. Trees can be blue, skies could be green, the ground could be on fire; decisions on what lives in the world alongside characters will be a character itself. This ultimately was part of the artistic considerations used in the voice over narrative, specifically "*Is the depiction of reality important in VR? How much reality is enough?"* It is in this question that artists must consider the purpose of reality in virtual reality and what limits can be pushed and why. It is in this way that volumetric video is key in the future of virtual reality because it does present an authentic signal of reality and a depiction of the real world brought into a virtual space.

SINGLE CAMERA VOLUMETRIC VIDEO

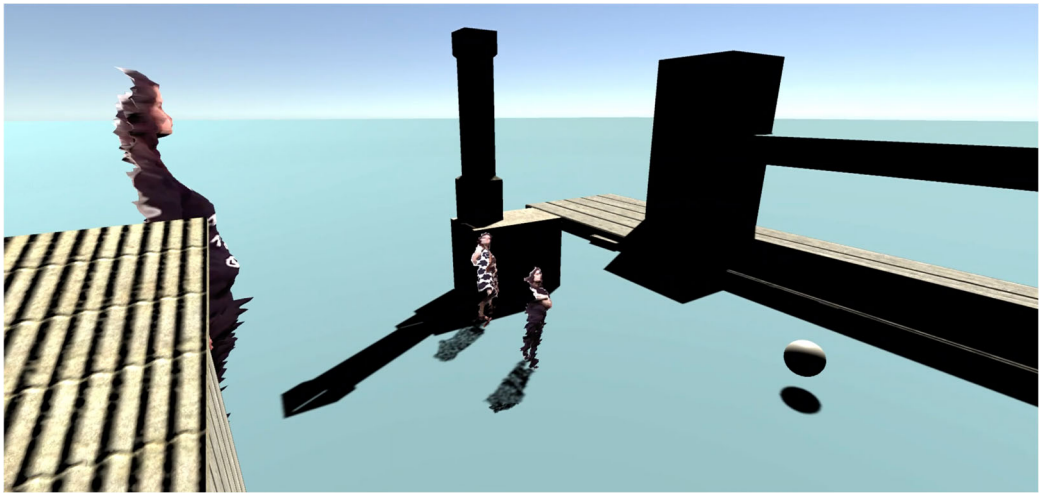

Image 18

When bringing in single camera volumetric video segments, we can see the video is fragmented and has glitches. We can see partial bodies moving in space. We know these bodies are digitally interpreted. Viewers intuitively understand volumetric video is filmed in the real world and placed into virtual reality, but it is clear we are not in a replicated real-world environment. The figures feel like they are from a different time with a message from the past or future. The same would be true in a photorealist virtual world context. As a viewer we do not have the same opportunity to speak to the characters and get feedback, as we would, for example, in Mozilla Spoke Hubs in a

social context. The volumetric characters do not engage with the gaze of the viewer and are omnipresent in the virtual world.

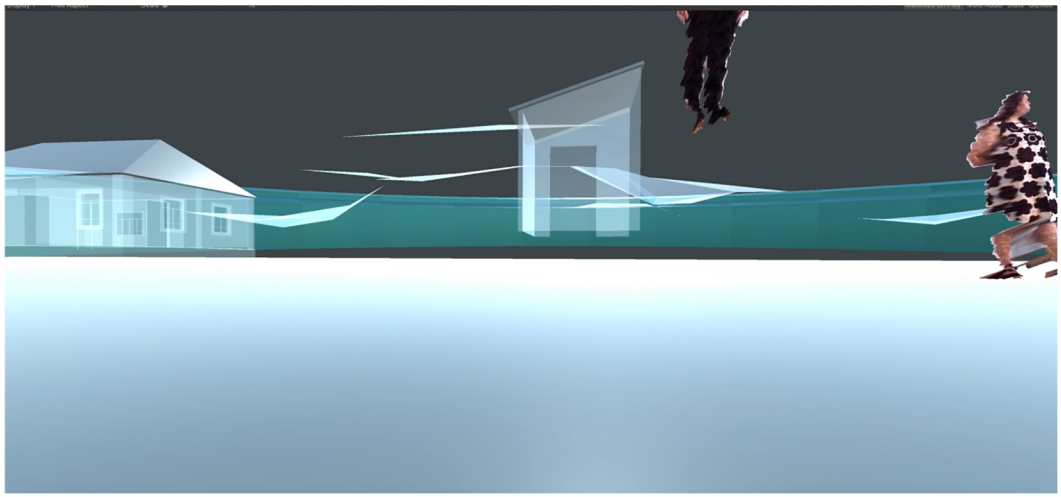

Image 19

We can consider how these bodies move in the world. They can be alone, or in a group. They can be larger than buildings floating in the air. Scale and gravity are choices that we can consider. How this is different than architecture or objects relates to materiality. Just as volumetric video cannot cover skeletons, as discussed, in the Radical web software research, material skins cannot cover volumetric video, as the video itself acts as a live skin that is mapped onto a geometric object. It is not a perfect representation and lacks the suspension of disbelief we have when we are looking at a live figure. It is in that understanding that we know we are in a fiction and telling a story. This insight contributed to several of the questions posed in the voiceover as outlined in the introduction that artists should consider when developing volumetric video for VR. Specifically*," How do the tools we use reflect the stories we tell? How do the depictions of bodies in 3-D shape the narrative?* and *how do we show history in VR?"*

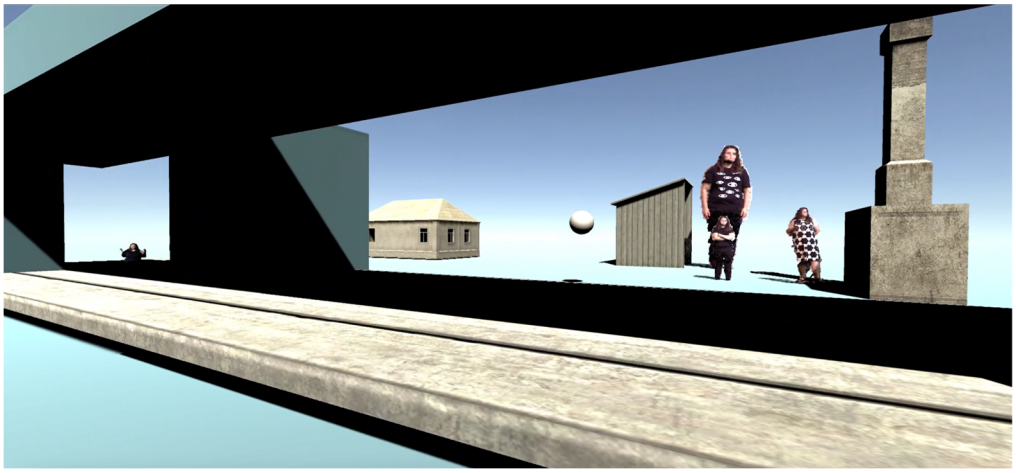

Image 20

VR CAMERA

Not only do we need to consider the depth camera when filming volumetric video, we also need to consider the camera providing the point of view for the viewer in the virtual environment. Like a film set, a camera can be placed on a rail path to create a dolly camera effect. The scene can be revealed over time, allowing the viewer to shift their point of view by looking around in a 360-degree capacity. This is a directed form of interaction where viewers are not permitted to freely roam spaces or need to complete a task.

As seen earlier in Image 1, the camera functions similarly to a film camera on a dolly, although it can also more closely parallel a camera drone. The path doesn't have to stay on the ground but can be placed in the air with more of a bird's eye view of the world.

Image 21

In building a point of view, we also need to consider the navigation. Does the architecture have hard surfaces or is it transparent where viewers can go through walls? Speed is also a factor. Too fast, and motion sickness can be a problem when designing an interaction on rails. Depending on the speed, the interaction could feel more like a rollercoaster versus a slow track into an environment. Not only do we have to consider the pace of the loops of the figures, but the pace of the interactive experience. Time is an important consideration when creating an immersive environment.

Image 22

SOUND

Sound is also something that needs to be considered and was critical to my own experimentation. Sound can be displaced and embedded in environments or objects, or function as a voiceover throughout the experience. It is also important to consider if the characters themselves use their voice, either spoken or as an internal thought. The sound could be instrumental, electronic, or even silent. Sound was the last element to be added to the experience. Throughout the development process the work was silent and tested in silence. A voiceover by the author about questions for considerations in developing volumetric video VR experiences as outlined in the introduction was added. This acts as an aural guide for the viewer as they navigate the VR experience.

ERRORS

Part of the development process is negotiating the errors. When working with plugins and versioned Unity software there will be conflicts. Part of an artist's approach to "hacking" commercial gaming software to create artworks is intuitive coding. By this, I mean to work with the errors and modify the approach as a way of artistic problem solving. Working in an iterative way to problem solve, and at times even recreate the work, was part of the research creation process. This research project has started over several times due to technical and author error. Screen shots and video walkthroughs were used to document the progress of the volumetric video VR and the creative problem-solving process.

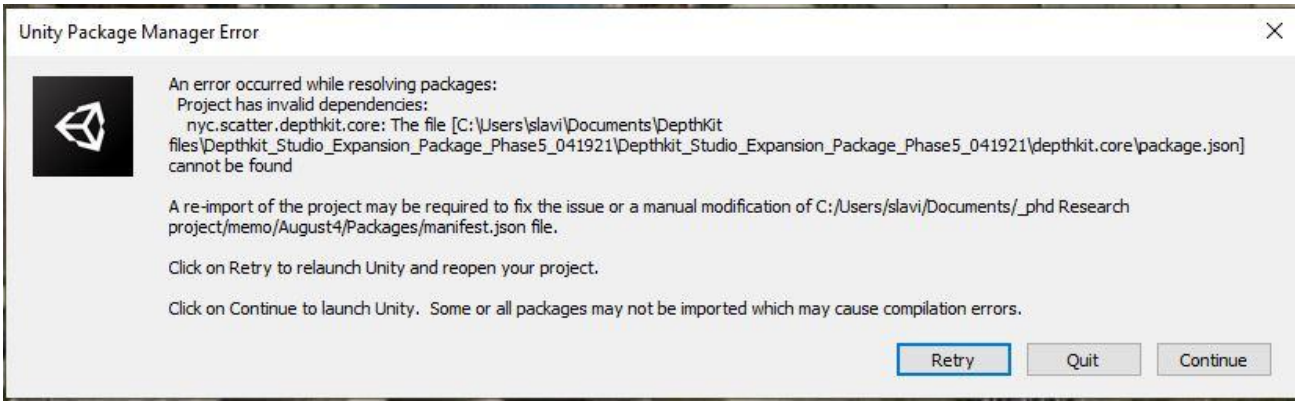

Image 23

4. CONCLUSION

The methodology of research creation was central to thinking through and testing ideas as I sought to understand the potential of volumetric video in the context of virtual reality and beyond. The volumetric studio of my living room became a laboratory both to test theoretical insights suggested in the literature and, in turn, to develop novel insights of my own. I also turned to other artists and theorists as I was working in isolation seeking interlocutors to these practices – who could help me to answer new questions I had around the gaze? The use of myself as model? Ultimately, I see my contribution to the field of volumetric scholarship through this research creation process as threefold. The first was to expand on Marks' definition of haptic cinema not only to include the embodied viewership of the viewer and image, but also the examination of the relationship of the author's self-gaze in the creation of an image of self for display.

The second was to expand on Papagiannis's definition of AR and VR and situate volumetric video. I argue that volumetric video in VR situates itself between Papagiannis' third convention, "ghosts", and the fourth - "living pictures". Therefore, volumetric video VR could be considered a subset of AR storytelling within a VR environment.

The third contribution to the field would be the questions developed through hundreds of hours of experimentation and iterative research creation for artists to consider when developing volumetric video for VR. The questions are a key provocation – grounded, and suggestive rather than prescriptive. This is useful to the field as it prompts meaningful considerations in the research-creation process to allow artists to think through concepts of volumetric video and contribute to the future of both volumetric form and the environments in which volumetric video may be embedded – specifically virtual reality in the case of this dissertation, but with implications for a range of XR environments.

The future of virtual reality will always be speculative given new technological innovations. My next steps will be to revisit the research of telexistence and web VR to

also include mobile volumetric AR experience. I would be interested in exploring how locative volumetric video AR experiences change and evolve storytelling interaction models and user experiences in a meaningful way. I am also interested in seeing the potential of live streamed volumetric video, exploring how it can be leveraged for social interactions in the future, how this will impact new ways of creation and collaboration for artists and viewers, and what new guiding questions might be suggested through experimentation with tools that support social volumetric work.

WORKS CITED

Back, Bettina. “Virtual Reality: How does the viewer get into the picture? A short reception aesthetic history of shifting frames.” *The Unframed World: Virtual Reality as Artistic Medium*; edited by Christoph Merian, 16-25. Verlag, Haus der elektronischen Künste, Basel, 2017 .

Bolter, Jay and Engberg, Maria. “Mobile Cinematics” Compact *cinematics: The moving image in the age of bit-sized media*. edited by Hesselberth, Pepita, and Maria Poulaki, 168-172 Bloomsbury Publishing USA, 2017.

Bolter, Jay and Gromala, Diane. *Windows and Mirrors, Interaction Design, Digital Ar and the Myth of Transparency.* Cambridge, MA: MIT Press, 2003.

Frayling, Christopher. “Research in Art and Design.” *Royal College of Art Research Papers*, 1, no. 1, 1993: 1-9

Friedberg, Anne. *The Virtual Window.* Cambridge, MA: MIT Press, 2009.

James, Valencia; Eyebeam. “Devising Performance for Immersive Webspace: Volumetric Performance Toolbox” Video Pool Lecture, online. February 2, 2021.

Jenkins, Henry. “Game Design as Narrative Architecture.” *First Person: New Media as Story, Performance, and Game.* Edited by Noah Wardrip-Fruin, and Pat Harrigan, 118-120. Cambridge, MA: MIT Press, 2004.

Lanier, Jaron. *Dawn of the New Everything: Encounters with Reality and Virtual Reality*. New York: Henry Holt, and Co., 2017.

Laurel, Brenda. “The Six Elements and the Causal Relations Among them.” *The New Media Reader*. Edited by Noah Wardrip-Fruin and Nick Montfort, 564-673. Cambridge, MA: MIT Press, 2003.

Lichty, Patrick. “The Cassandra Complex: An oracle of Virtual Reality Art.” *The Unframed World: Virtual Reality as Artistic Medium*. Edited by Christoph Merian, 46-57. Verlag, Haus der elektronischen Kunste, Basel, 2017.

Loveless, Natalie. *How to make art at the end of the world.* Durham, NC: Duke University Press, 2019.

Marks, Laura U. *The Skin of the Film Intercultural Cinema, Embodiment, and the Senses*. Durham, NC: Duke University Press, 2000.

Papagiannis, Helen. *Augmented Human: How Technology Is Shaping the New Reality*. Sebastopol, CA: O'Reilly Media, 2017.

Roquet, Paul. “Telepresence Enclosure: VR, Remote Work, and the Privatization of Presence in a Shrinking Japan.” *Media Theory* 4, no. 1, November 2020: 33-62

Ryan, Marie-Laure. *Narrative as Virtual Reality*. Baltimore, MD: John Hopkins University Press, 2001.

Steyerl, Hito. "Ripping reality: Blind spots and wrecked data in 3D." *Website of the EIPCP (European Institute for Progressive Cultural Politics)* (2012).

APPENDIX

The following is documentation of artist talks and a video walkthrough of the research creation process.

"Documentation – Considerations for Volumetric Video in VR." *YouTube,* uploaded by Slavica Ceperkovic, 28 April 2022, youtu.be/UULuU1ThyfA

"Research Creation Process." *YouTube,* uploaded by Slavica Ceperkovic, 10 June 2022, youtu.be/9musyS6BBGo

"Speaking Volumes." *YouTube*, uploaded by Technomentary Lab, OCAD University, 11 February 2022, youtube.com/watch?v=8e1t2bYd7GU. (51:49 – 1:07:10)

"STYLY x InterAccess XR Talk." *YouTube*, uploaded by InterAccess, 29 April 2022, youtube.com/watch?v=ybLpau4mSXU (36:50-1:01:40)