The Final Image: An Artwork and Techniques in Immersion

Martin Jarmick

A dissertation submitted in partial fulfillment of the requirements for the degree:

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Richard Karpen, Chair

Juan Pampin

James Coupe

Phillip Thurtle

Authorized Degree granting program:

Center for Digital Art & Experimental Media

University of Washington

Abstract

The Final Image: An Artwork and Techniques in Immersion

Chair of the Supervisory Committee, Richard Karpen

Center for Digital Art & Experimental Media

The Final Image is a virtual reality (VR) artwork combining spherical stereoscopic images, interactive interludes, 3D soundscape, and voice-over narration. Viewers use a head-mounted display (HMD) and a hand-held controller to navigate 3D spaces and activate portals into immersive narrative vignettes. The piece takes the form of a cinematic visual-sonic poem about fragmented recollections of home, and an inner journey of reassembly and clarity. The project is regularly updated and publicly available on the internet where it can be downloaded and viewed using the appropriate virtual reality hardware.

This document supports the production by describing a framework for narrative immersion in the artistic context. Works of art in cinema, painting, music, video installation, electronic games and literature are presented as influences for The Final Image, and in addition, media theory and philosophy lay a foundation for the phenomenological structure of immersion with both subjective and spatial dimensions. Discussions of artworks and concepts weave in the document and reflect the non-linear relationship of thought and practice present in The Final Image.

Table of Contents

List of Figures

Acknowledgements

First and foremost, I wish to acknowledge and express gratitude to Richard Karpen for his advice and mentorship over the years. His contribution to arts research is extraordinary and he inspires and challenges myself and my colleagues in the deepest of ways. I would also like to recognize James Coupe and Juan Pampin for supporting my work, and Phillip Thurtle for passing on to me his curiosity and passion for ideas and new knowledge.

Additional gratitude goes to my friends and colleagues who participated in this research: The cast and crew of *The Final Image*, Jack Straw Cultural Center, The Grand Illusion Cinema, Heather Raikes, Robert Twomey, Michael McCrea, Marcin Pączkowski, Ha Na Lee, James Hughes, Benjamin Mentzer, Daniel Peterson, Joseph Anderson, Yunmi Her, Chun Shao, Tivon Rice, Pete Moss, Jesse Blanchard, Cynthia Caci, Kevin Mihata, Afroditi Psarra, Donald Craig, Jeffery Frace, and Blake Hanniford. Further, a special mention to Andrew Theisen.

Finally, I would like to note appreciation for the professionals at Valve Corporation and Unity for their generosity of time and resources.

To Stephanie, Pearl and my parents. A dissertation is truly a family affair

Project Production Credits:

Andrew Theisen: Technical Artist / Programmer

Terence and Angela Kelley: Voice Talent / Actors

Adam Hogan: Camera / Imaging Associate

Michael McCrea: Programmer / Technical Assistance

Yun Mi Her, Sarah Maglaras: Assistance

Section 1: Introduction

1.1: Project Description & Context

The Final Image is a virtual reality (VR) artwork combining spherical stereoscopic images, interactive interludes, 3D soundscape, and voice-over narration. Viewers use a head-mounted display (HMD) and a hand-held controller to navigate synthetic 3D spaces and activate portals into immersive narrative vignettes. The piece takes the form of a visual-sonic poem about fragmented recollections of home, and an inner journey of reassembly and clarity. The production was executed over the course of two years and involved writing, actors, cinematography, image-processing, three-dimensional modeling and animation, sound design, user-experience design, and software engineering. *The Final Image* project is regularly updated and publicly available on the internet where it can be downloaded and viewed using the appropriate virtual reality hardware.

This document supports the production by describing a framework for narrative immersion in the artistic context. Works of art in cinema, painting, music, video installation, electronic games and literature are presented as influences for *The Final Image* and discussed to define an epistemology. In addition, media theory and philosophy lay a foundation for the phenomenological structure of immersion in the subjective and spatial dimensions. Discussions of artworks and concepts weave in the document and reflect the non-linear relationship of thought and practice present in *The Final Image.*

The purpose of a PhD in any field is the allowance of time and space resources to conduct and contribute original research and define a trajectory for further study and career. The criteria for successful research at the PhD level varies with institution and discipline, but there is an agreed level of thoroughness and discovery necessary to achieve success. This spectrum is agreed upon through deliberation of committees, consisting of internal and external members. These general features of doctoral study are present here, and in part, govern the requirements of this paper and project. And like the much of the lore of discovery both scientific and artistic, inspiration comes at such instances as a walk down the street at night, or hitting one's head, or hearing a musical melody, or watching an apple drop from a tree. At DXARTS, graduate students and faculty are examining works of art, past and present, and conducting experiments to find research questions that combine personal inspirations and relevance in the field at large. The purpose of this paper is to summarize and archive the experiments and results of this project in hopes that further inquiries by future individuals can be established and accelerated.

1.2: Terms

It is necessary to clarify and orient several terms in this document to avoid confusion and position the research context. First the term *cinema* or *cinematic* describes a set of features that define form such as a shot, a montage, and sound, and when these variables are applied to a subject, a space-time relationship emerges forming a cinematic moment. Cinema is a language, a framework for creative expression that gelled in its specific mechanical and socio-economic condition. This document avoids examples in

entertainment-oriented production and primarily focuses on the art cinema tradition as defined by pioneers such as Carl Theodor Dreyer, F.W. Murnau, Dziga Vertov, Fritz Lang and Sergei Eisenstein to name a few, who leaned toward the emotive, haptic, and poetic dimensions of cinema. Of course, work by these and current cinema artists are in proximity to entertainment and the defining lines can be complex and unclear. For instance, this document cites important cinematographic contributions by Gregg Toland, who often worked in studio feature production on important films such as Citizen Kane and The Grapes of Wrath. As early film theorist Rudolf Arnheim stated in 1933, "Film resembles painting, music, literature, and the dance in this respect – it is a medium that may, but need not, be used to produce artistic results."[1]

Similarly, the term *game* is used to describe several works in this document, yet these works do not adhere to popular conventions of the form. As in cinema, the platform is appropriated for entertainment and other knowledge areas such as education and scientific research. In this research, a game is defined as a navigable experience or responsive environment mediated by viewer input control, and the pieces cited, extend storytelling and poetic techniques. Tracy Fullerton, artist and game designer of several open-world and interactive narrative works, cites the *art game* as an emerging genre noting that "Some people would not call these examples games, but as the commercial market for games diversifies, I believe that it is communities of innovation such as these that will drive new ideas and point new ways to new forms of play and interactivity."[2]

[1] Arnheim, Rudolf. *Film as Art*. ACLS Humanities E-Book. Berkeley: University of California Press, 1957.

[2] Fullerton, Tracy. *Game Design Workshop: A Playcentric Approach to Creating Innovative Games, Fourth Edition*. 4th ed. Milton: Chapman and Hall/CRC, 2018.

It is necessary to define *experimental* in order to dispel a common use of the word as a genre of production, and reaffirm its original definition maintained mostly in the scientific context. An experiment, in broad terms, is part of a methodology. It is an action put forth to discover new knowledge. For instance, a series of experiments are iterated to obtain results, which are analyzed in preparation for further experiments or used to prove theoretical aims. When we hear the terms experimental art, experimental music, or experimental cinema, this has inherited a qualitative property where for instance, an experimental film uses non-linear narrative technique, non-diegetic relationships or heavy degrees of image abstraction as opposed to conservative or linear technique in much entertainment media. Laura Marks' articulation on experimental moving-image production is useful in this discussion:

> It includes films and videos that experiment formally with the medium, from film formats to low-end video formats to HD to mobile and online platforms. It includes experiments drawn from critique of cinema and TV, with sound, montage, structure reflexivity, etc. It experiments with the relationship between fiction and documentary: presence, index, performance…Experimentation also regards content: experimental narrative, essay film, experimental documentary…[3]

The Final Image is experimental in several ways. For instance, it proposes that immersive narratives must be methodically porous and open to enable audience participation in the poetic image presented. Next, the project blurs a binary between an active and passive media experience with fixed and interactive structures operating in

[3] Marks, Laura U. *Immersed in the Single Channel: Experimental Media from Theater to Gallery*. *Millennium Film Journal* no. 55 (2012): 14-23.

parallel. Further, the work adapts cinematic techniques to an emerging platform in order to discover new knowledge about the language such as the agency-environment standpoint discussed in this research.

1.3: Structure

This document has four main areas including this first introductory segment to summarize the project and establish terms, context, and structure. Section 2: *Project Framework* is concerned with the works of art, theory, and ideas that inform and position the research and consists of four parts. First, *Cinematic Immersion* illustrates ways in which screen-based cinematic works of art immerse audiences through two-dimensional means. Early theoretical ideas regarding liberated screen-space and viewer choice are cited to commence the discussion of immersion, and cinematic works from the early twentieth century though the present are discussed. The second section, *Agency-Environment,* shifts the inquiry into ways in which painted works and contemporary video art allow viewers to inhabit and navigate the pictorial field. Artists such as Bill Viola use cinematic moving-image techniques to create contemplative moments in physical space, extending a definition of immersion. *Interaction & Navigation* discusses metaphysical and literal views on interaction and responsive environments in electronic games, establishing phenomenology as a useful lens to conceptualize the medium's correlative nature. *In Literature, In Virtual Reality*, discusses approaches to perspective and ambiguity that expand the literary image. In addition, *The Final Image* is introduced in the literary, conceptual, and formal terms established in the chapter.

Section 3: *The Final Image* organizes around the research questions and production areas of the piece itself. The first part, *Research Objectives,* summarizes the research problems and how they are addressed, and provides context for the content in the work. The second part, *Preceding Works*, illustrates a continuum of research and a trajectory of techniques and project work that led to *The Final Image*. The subsequent segments, *Writing and Voice Recording, Sound Design and Composition, Location Photography, System Overview*, and *Modeling and Animation* detail the techniques implemented in the project and reflect a manifestation of the ideas in the *Project Framework*. Diagrams, technical images, production stills, and code samples act as visual aids for the descriptions. Specific product names of software tools and image / audio capture are omitted to prioritize the underlying processes and techniques and to prolong obsolescence as technology exponentially changes.

Section 4: *Conclusion / Preclusion* summarizes the document and project and describes the immediate and long-term effects on a research trajectory moving forward. This includes new online distribution platforms for *The Final Image*, submission to emerging technology categories at film festivals, and an upcoming exhibition of the work in an art space.

Section 2: Project Framework

The image is indivisible and elusive, dependent upon our consciousness and on the real world which it seeks to embody

Andrei Tarkovsky, Sculpting in Time

2.1: Cinematic Immersion

The term cinematic is ubiquitous and has come to describe a perceptual phenomenon regardless of medium. For over a hundred years, artists and engineers have subverted limitations of the medium to develop spatial and subjective dimensions, ultimately forming a language system and discovering ways to immerse viewers within the pictorial field. For instance, Carl Theodore Dreyer's film, *The Passion of Joan of Arc (1928),* avoids the literal depiction of architectural space and is primarily made up of protagonist close-ups allowing the audience to inhabit the attention and emotions of the main character (Figure 1).

Figure 1: Various instances of look directions from close-ups in the film

Establishing shots of physical spaces are fleeting or do not exist, so most of the spatial relations are established through eye-lines. In one sequence, the central character Jeanne D'Arc's transfer from a trial room to an interrogation chamber is expressed by a close-up of her feet stepping over a small barrier, followed by a montage of torture devices, then back to close angles of her face. Dreyer creates a prudent expression of space and time, enabling the audience to form schema from very little information.

An immersive media experience is the fusion of agency and stage. In other words, immersion is a circuit between active audience perception and environment. Dreyer's film succeeds in enveloping his viewer in a visceral and direct way by positioning the them at close proximity to the subjects. As theorist, Rudolf Arnheim observes, "Dreyer in his Passion of Joan of Arc stresses a monk suddenly jumping up excitedly from his seat by placing the camera closely in front of the actor so that through this forward movement, his figure grows to an enormous size and occupies the whole screen."[4] (Figure 2)

Figure 2: The monk sitting (left), then suddenly standing into a low angle and subordinating the audience

Although traditional forward-facing modes of viewing are considered passive, the dark space and relatively large projection surfaces in theatrical viewing allows audiences to

[4] Arnheim, 61

enter screen universes. Viewer navigation in two-dimensional cinema space is aided in part by allowing access to as much of the field as possible. For Andre Bazin, author and pioneering theorist, audience choice and participation in the picture is increased with deep depth of field and long takes and a "dialectical step forward in the history of film language."[5] Using the cinematographer Gregg Toland's work to illustrate, he continues:

> …it implies a more active mental attitude on the part of the spectator and a more positive contribution on his part to the action in progress. While analytical montage only calls for him to follow [the director's] guide, to let his attention follow along smoothly with that of the director who will choose what he should see, here he is called upon to exercise at least a minimum of personal choice. It is from his attention and his will that the meaning of the image in part derives.[6]

Audience attention is guided by the director's use of light, motion, and scale, and at the same time, the viewer is free to roam in Toland's deep visual fields (Figure 3).

Figure 3: Cinematographer Gregg Toland. Left: The Westerner (1940), right: Citizen Kane (1941)

[5] Bazin, André, Andrew, Dudley, Gray, Hugh, and Renoir, Jean. *What Is Cinema?. Volume 1*. Berkeley: University of California Press, 2005, 35.
[6] Ibid, 36

Similarly, we see this balance in the work of Andrei Tarkovsky where viewer immersion in the cinematic image is enabled with deep focus, rich visual fields, and extended durations. For instance, in the final shot of his film *Nostalghia (1983),* the camera starts on the protagonist in full, then reverse dollies into an extreme wide shot over the course of nearly three minutes (Figure 4).

Figure 4: Shot from Nostalghia starts in full (top), then ends in wide (bottom) over course of nearly three-minutes

Similarly, for Tarkovsky the essential cinematic moment exists in the moving-image itself and "…flows beyond the edges of the frame"[7], regardless of montage. He heeds viewer attention and immersion with elongated shots, unending focal distance, and periodically manipulates film-speed to expose time even further. These devices combined with compelling subject and pictorial composition push cinema beyond its limits:

> The film then becomes something beyond its ostensible existence as an exposed and edited roll of film, a story, a plot. Once in contact with the individual who sees it, it separates from its author, starts to live its own life, undergoes changes of form and meaning.[8]

Extended takes are a primary technique in all of Tarkovsky's films and are seen with varying degrees of movement, blocking, and focal areas. For instance, in his film *Mirror (1975)*, the camera dollies, pans, and tracks through a home following the characters as they run outside to see a fire. The shot is several framings in one motion resembling a montage without actual edits. Similarly, the artist Bela Tarr creates continuous cinematic moments with orchestrated sweeps of scene spaces. In his film *The Man from London (2007)*, Tarr surveys a shipping port, positioning his audience in a watchtower for a panoramic observation of the space. The camera slowly follows various actions of a crime being committed around the tower in a twelve-and-a-half-minute shot (Figure 5).

[7] Tarkovsky Andrei. *Sculpting in Time: Reflections on the Cinema*. London: Bodley Head, 1986, 118.

[8] Ibid, 118

Figure 5: Still images depict eight camera positions from the twelve and a half minute continuous shot

As author and critic Manohla Dargis observed of the work, "There are moments when watching one of Tarr's films that it seems as if he doesn't just want you to look at his images, but to somehow enter into them alongside the characters."[9] The director's deep

[9] Dargis, Manohla. "And Now for Something Completely Meditative." *The New York Times,*2007, Late (East Coast).

depth and duration enables audience to settle in, develop agency, and inhabit the cinematic world as he unfolds it with light, scale, and movement.

2.2: Agency-Environment

Historically, artistic uses of agency and the visual field for immersion dates back to the Paleolithic illustrations rendered in the caves of Southwest France over ten-thousand years ago. In his article, The First Picture Show: Cinematic Aspects of Cave Art[10], Edward Wachtel describes etched and painted forms depicted on irregular rock surfaces deep in underground caverns. Superimposed line drawings become moving images when viewers move light across the cave surface at particular rates and positions, communicating and storing information for hunting. Movement of the forms is interactive, dependent on viewer control of the light source, and the dark and quiet characteristic of the cave spaces provides a theater where young hunters can study animated images of animal behaviors. Moving ahead a few thousand years, Michelangelo's Sistine Chapel Ceiling exhibits a visual narrative spanning over one-hundred and thirty feet and utilized advancements in painting technology to augment its dimension and vibrancy. The artist extended the one-point perspective technique popular at the time by creating multiple vanishing points along the length of the mural (Figure 6). This liberates the focus for a viewer at any given position making the picture an open and navigable field.

---

[10] Wachtel, Edward. *The First Picture Show: Cinematic Aspects of Cave Art*. *Leonardo*26, no. 2 (1993): 135-40.

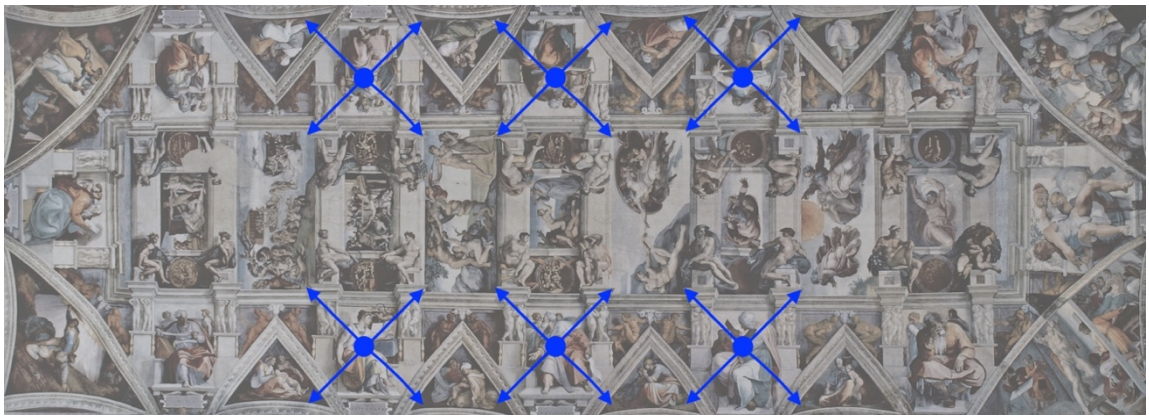

Figure 6: Blue arrows illustrate multiple vanishing points along the Sistine Chapel Ceiling

In the same Renaissance period, the artist Hieronymus Bosch created a large polyptych painting called The Garden of Earthly Delights () that engulfs viewers in a three-act narrative (Figure 7).

Figure 7: Bosch tryptic panels express a continuous narrative field of activity

Bosch's work adapts freely to audience position with more oblique view-points than the Renaissance perspective techniques of the time. Additionally, the immersive visual environment is emphasized where the separate panels are divided structurally, yet continuous in content and theme. Portals such as doors, pools, and holes connect the

three worlds, and in one case, mutant creatures from the dark third panel climb ashore into the pristine light of the first panel expressing a cycle between the opposite realms. The immersive agency-environment relationship is further explored in the work of Caspar David Friedrich where he enables audiences to inhabit his landscapes through the use of human foreground characters or the Rückenfigur (Figure 8). As art historian Joseph Koerner notes regarding the figure in Friedrich's Wanderer Above the Sea Fog:

> It is he who mediates our experience of the scene, and who knits together the landscape's disparate fragments. Indeed, it is hard to imagine what the view from the summit would be without his centralizing and concealing presence, how, for example, the symmetrical hills radiating from just below his shoulders would actually meet in the valley. The *Rückenfigur* is so prominent in the composition that the world appears to be an emanation from his gaze, or more precisely, from his heart.[11]

The observer characters frequent in Friedrich's paintings, push his landscapes beyond a regarded object, and instead, frame them as a perceived and emotive space. Further, even in the absence of a Rückenfigur, his works engulf viewers with sites whose geological form, natural objects, and archeological relics are extensions of the mind and archetypes connected to consciousness. Koerner continues:

> Even without an internal viewer in the picture, Friedrich's landscapes present themselves as something seen, rather than simply as something there. Their

---

[11] Koerner, Joseph Leo. *Caspar David Friedrich and the Subject of Landscape*. 2nd ed. London: Reaktion, 2009, 213.

symmetrical design suggests the presence of an eye arranging nature, and the *Rückenfigur* only advertises this presence. In a sense, he is redundant. Is not our eye enough, coupled with the evidence of the artist's arranging eye, to infuse the landscape with a beholder's gaze, or do we need another hidden look, stationed within the painted world?[12]

Figure 8: The observer-character is a vessel in which the viewer inhabits the scene.

[12] Ibid, 213

Dramatic narrative and high-fidelity form and texture are characteristic of works in the Renaissance and Romantic periods, and the aforementioned techniques in literal and figurative perspective served to unify the viewer and picture.

Artists working in various aural and visual arts have utilized and extended cinema grammar and technology to create moving-image experiences outside of conventional production and theatrical release. To contemporary artist Bill Viola, picture scale, movement, and image fidelity is a doorway into his work, compelling viewers to gaze and inspect. In addition, his works take a comprehensive approach to creating an immersive experience stating that, "Experience is so much richer than light falling on your retina. You embody a microcosm of reality when you walk down the street – your memories, your varying degrees of awareness of what's going on around you… representing that information is going to be the main issue in the years ahead…"[13]. His work with video is exemplified in The Crossing (1996) which uses a freestanding, double-sided projection surface positioned in an open exhibition area surrounded by four speakers. On both screens, a human figure slowly walks toward the viewer then pauses. One figure is then consumed by a torrent of falling water, and the other, the same figure erupts into flames (Figure 9).

[13] Rutledge, Virginia. "Art at the End of the Optical Age." *Art in America* 86, no. 3 (1998): 443-452.

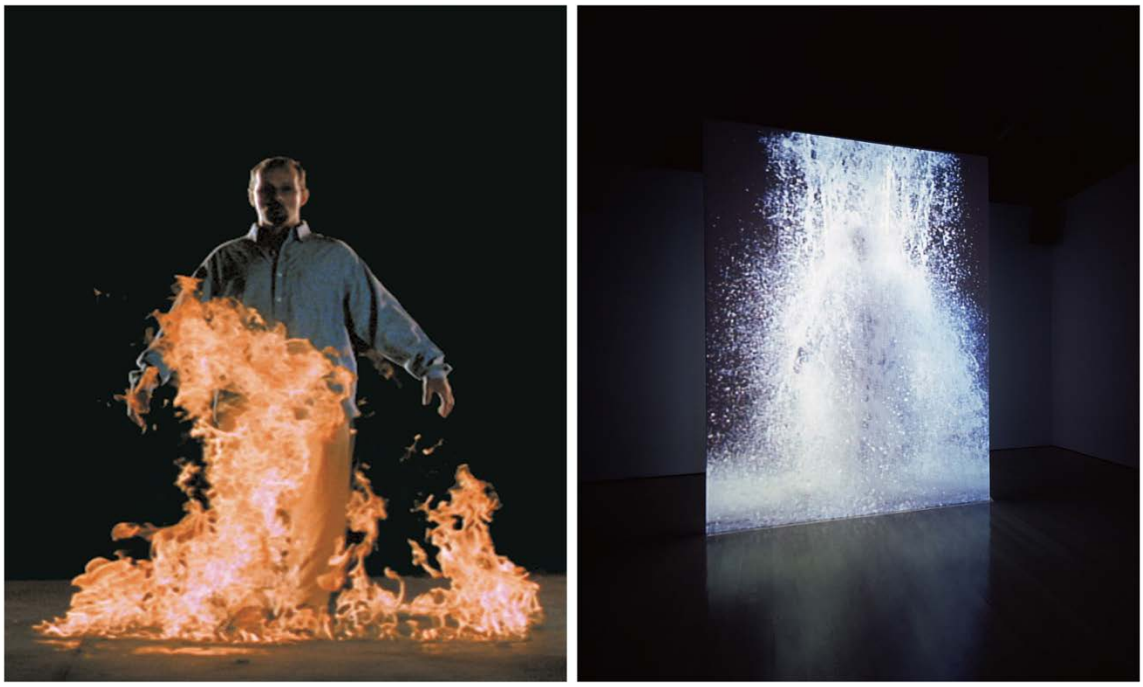

Figure 9: Exhibition stills of scenes of water and fire engulfing the body

The viewer is primarily engulfed in darkness, confronted by the nearly ten-foot tall moving images centered in the space with surrounding sound. Navigation is encouraged by the limited visibility of both moving images – the viewer is compelled to reposition themselves to see each side of the projection surface. Viola creates a hypnotic condition for the audience by using significantly slow frame-speed with continuous movement, elemental fire and water, and isolating the audience senses with dark space. Similarly, Gary Hill's Tall Ships (1992) presents a navigable moving-image theater where audiences walk down a long dark corridor with sixteen projections of people on either side (Figure 10).

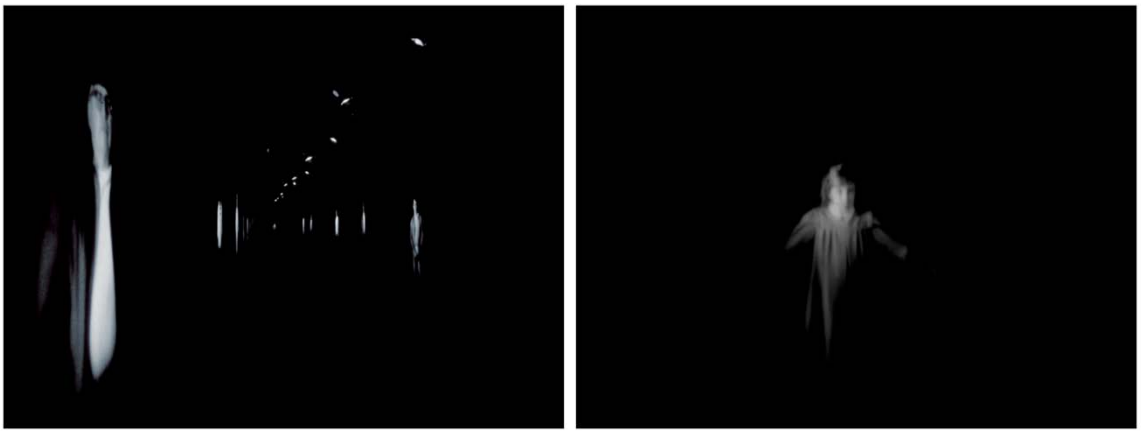

Figure 10: Corridor of projections (left), close up of approaching character (right).

In addition, Tall Ships uses viewer position to activate the work with sensor technology integrated into the video playback system. In relation to this correlative aspect in the piece, Hill explains, "I wanted the interactivity to be virtually transparent…I didn't want somebody going in there and, in a sense, playing."[14] In the initial state the figures are distant, and as viewers approach a projected image, the figures come forward. The subtle interactions create an illusion of living beings in the sixteen orchestrated encounters with viewers. The artist continues, "The feedback makes possible a special state that engages an ontological space where art and life mutually fold into each other."[15]

2.3: Interaction and Navigation

Most art can be considered interactive. Even in the absence of a literal input/output circuit, an artwork unfolds to its audience, the audience becomes familiar with the world that's presented, and the artwork subsequently spawns new meaning. The artist anticipates this temporal and cyclical process by designing possibilities and complexity

[14] Quasha, George, Hill, Gary, Stein, Charles, and Cornwell, Regina. *Tall Ships*. Gary Hill's Projective Installations; No. 2. Barrytown, New York: Station Hill Arts, 1997, 440

[15] Ibid, 441

and, in this way, a production is never encountered with total indifference. Theorist Laura Marks challenges the position of passivity in cinematic viewing stating:

> 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. Rather than witnessing cinema as through a frame, window, or mirror, the viewer shares and performs cinematic space…[16]

This metaphysical view of interaction is not trivial or incidental: artists of all mediums build a relationship with their viewer by crafting an information exchange with the senses. For instance, the Impressionist painters used new science in human optics to more deeply integrate their images with viewer perception, and composers throughout history use harmonious and discordant relationships with aural biology to move their audience. In both cases, artists are generating patterns that transmit to viewer receptors, which feedback into the artwork as a system of perceptual possibilities.

Electronic games exemplify a literal expression of interaction in their presentation of responsive environments and underlying system dynamics: the viewer and content are necessarily correlated in a direct way. Game designers and artists conceive narratives, conflicts, and tasks then compose visual and sonic fields that respond to audience input

---

[16] Marks, Laura U. *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*. E-Duke Books Scholarly Collection. Durham: Duke University Press, 2000.

with varying degrees of freedom. In Dear Esther (2012), an open-world exploration piece created in the first-person perspective, the audience uncovers a story by discovering fragments of narration, sound, and objects embedded in the landscape of a mysterious island (Figure 11).

Figure 11: Four areas from the island in Dear Esther

The world of Dear Esther is a three-dimensional canvas conceived with themes of loss, love, redemption, and secrets, and similar to Friedrich's landscapes, the topology is arranged to enable audience curiosity and emotional experience (Figure). Attention, direction, and navigation is guided via paths and openings carved in the terrain and visual beacons such as lighthouses, shipwrecks, and curious candle memorials. In addition, the sonic world unfolds in relation to viewer location, and as composer for the work Jessica Curry acknowledges about the production process, "Music design and narrative formed a very deep fusion, and no art form took precedence over the other… music wasn't the

slave to the visuals, and instead, formed an integral part of the experience."[17] In this way, the work diverges from conventional production and design modes common in electronic games, and instead favors composition methods similar to music or art cinema. The dramatic function of the voice-over text segments and sonic materials is maintained no matter what order they are enacted by the audience. In a sense, Dear Esther is an immersive narrative system mobilized through the relationship of viewer choice and environment. It is a passive / active hybrid: duration and order are dynamic, yet the authored thematic integrity is intact.

An embodied vision, defined by a plurality of senses, consciousness, and mobility, is a phenomenological sight, and a useful philosophical frame to consider the electronic game medium. The vision defined by theorist Maurice Merleau-Ponty describes our environment as opening its nature to our senses, an interchange between our surroundings and our subjective position, observing that, "I do not see space according to its exterior envelope; I live it from the inside. I am immersed in it. After all, the world is around me, not in front of me."[18] This rejects a Cartesian conception of experience where the body is divided from and objectively observes the space it occupies. Instead, Merleau-Ponty situates perception as relational, stating that "Since things and my body are made of the same stuff, vision must come about in them…light, color, depth, which are there before us, are there only because they awaken an echo in our bodies..."[19] Using painting to illustrate, he describes a circuit: a sensed object in the world reverberates an internal

---

[17] Curry, Jessica. *The Music of Dear Esther: Creating Powerful Scores with Limited Resources.* Conference Talk. Game Developers Conference (GDC), San Francisco, CA, 2013.

[18] Merleau-Ponty, Maurice, and Edie, James M. *The Primacy of Perception: And Other Essays on Phenomenological Psychology, the Philosophy of Art, History, and Politics*, 178.

[19] Ibid, 164

equivalent, and this correspondence is depicted on the canvas. An encounter with a painting should enable the viewer to inhabit the world or as Merleau-Ponty says, "Rather than seeing it, I see according to, or with it."[20]

Although philosophy often lives in the luxury of pure conception, phenomenology provides a platform for analysis and architecture in composing dynamic media experiences. Gone Home (2013) is a narrative world composed of three-dimensional spaces, objects, and sounds. The work unfolds as the viewer, or player, assumes the identity of a daughter returning home to find her family has disappeared. This contrasts the ambiguous protagonist inhabited by the audience in Dear Esther, and instead, maps a specified dramatic point-of-view or Rückenfigur into the world. The audience navigates the abandoned home, uncovering information about past and present residents by opening drawers, cabinets, and looking in crannies to find written notes, photos, and audio tapes (Figure 12). Select objects and entryways in the space are interactable allowing the viewer to pass through doors or pick items up for examination and inventory addition. The work diverges from Dear Esther in story, but more importantly, in structure. For instance, order is imposed on the game framework in Gone Home, where specific objects and spaces must be discovered in order to access other areas of the home. In a sense, this fixedness narrows viewer freedom yet facilitates strategy and cognitive-emotional discourse. Consequently, this spatiotemporal lattice in the form of ordered events and conditions ensures compositional cohesion designed by the author. In other words, the rules help propel a narrative direction and prevent the audience from straying too far.

[20] Ibid, 164

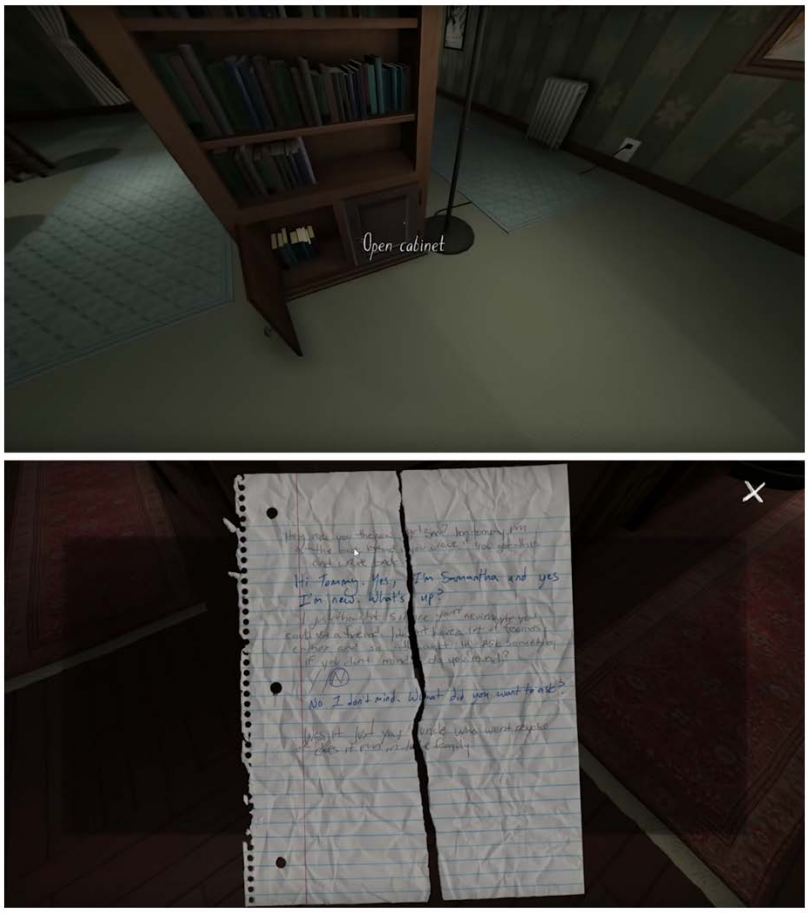

Figure 12: A prompt to open a cabinet (top), and a close up of a note (bottom) in Gone Home.

The delimitations or structures that exist in space are theorized as 'the striated and the smooth' in Gilles Deleuze and Felix Guattari's seminal phenomenological study, A Thousand Plateaus. They use composer Pierre Boulez's artistic methods to illustrate the framework stating that, "In the simplest terms, Boulez says that in a smooth space-time one occupies without counting, whereas in a striated space-time one counts in order to occupy."[21] The striated-smooth configuration can be applied to most artistic production through a relation between measured and fluid properties. In cinema, formal limitations such as the frame, focal distance, and narrative-rational conventions are a grid in which

[21] Deleuze, Gilles, and Guattari, Félix. *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press, 1987, 477

the subtle, uncertain, and poetic movements of performance and exposure meander. As the camera follows the protagonist perform a candle ritual in Tarkovsky's Nostalghia, in the laws governing the cinematography and mise-en-scene, the moment manifests, allowing the audience to witness the character's trial, emotionally drawn into the gestures, the failures, and the flickering of the candle's flame. Deleuze and Guattari continue in relation to Boulez and music:

> …the striated is that which intertwines fixed and variable elements, produces an order and succession of distinct forms, and organizes horizontal melodic lines and vertical harmonic planes. The smooth is the continuous variation, continuous development of form; it is the fusion of harmony and melody in favor of the production of properly rhythmic values, the pure act of the drawing of a diagonal across the vertical and the horizontal.[22]

Both Dear Esther and Gone Home articulate a mixture of the two structures wherein the former is more 'nomadic' and the latter 'sedentary'. In addition, each offers a different approach to agency in how they identify the player-protagonist, or Rückenfigur. Author Colin/Ciara Cremin articulates a relationship of Deleuzian theory to electronic games, observing that the medium "crosses a boundary between the screen depiction and the player perception to create a tangible sense of copresence."[23] In addition, Cremin extends the striated-smooth correlation by defining the game world as a friction-image, citing physics embedded in surfaces and zones as a distinguishing characteristic of

---

[22] Ibid, 478

[23] Cremin, Colin/Ciara. *The Formal Qualities of the Video Game: An Exploration of Super Mario Galaxy with Gilles Deleuze. Games and Culture* 7, no. 1 (2012): 72-86.

navigation in virtual realms. These forces are crucial in an environment-viewer integration, for they are the point where one makes contact with the limits or frame of the volume-image and poetic material. Deleuze and Guattari continue, "Whereas in the striated forms organize a matter, in the smooth materials signal forces and serve as symptoms for them."[24]

Returning to the work of Bill Viola, his project The Night Journey (2007-18) is a navigable visual-sonic artwork developed for an electronic game console (Figure 13). The viewer uses input controls to move slow-motion through an obscure landscape consisting of canyons, forest, ocean, and desert. Similar to the reader in literature, the audience-protagonist is an unnamed seer roaming the world, a hybrid third / first-person and a vessel in which the environment opens itself to. Dwellings appear and disappear as mirages, and portals present themselves, leading through tunnels and superimposed videos. The visual aesthetic resembles a low fidelity film stock, with heavy grain, blurs, and motion distortions obscuring the image legibility, and the soundscape enhances the overall abstraction with stretches of pattering noise, periodic whispers, and drone. Each area holds a defining set of sound materials, and in addition, the audience can activate moving-image vignettes that overlay the landscapes to express a flashback narrative. In a sense, The Night Journey represents an inner sight, denying the audience a clear exterior world or plot structure, and instead, presents a friction-image with player mechanics that enable viewer reflection. Tracy Fullerton, The Night Journey collaborator and game designer on the project, states they designed the narrative as "…a resistance to specific

[24] Deleuze / Guattari, 479

storytelling, and instead, the creation of a visual oral framework. As if a story was there but wasn't being told."[25]

Figure 13: Moments from The Night Journey, combines grainy game spaces with video vignettes.

[25] Fullerton, Tracy. *Breaking the Mold: Experiments in Evolving Game Narrative from the USC Game Innovation Lab.* Conference Talk. Game Developers Conference (GDC), San Francisco, CA, 2014.

2.4: In Literature, In Virtual Reality

Agency and inhabitance in a literary image are achieved partly by enabling reader participation and manipulating perspective, and The Night Journey, Dear Esther, and Gone Home use these to varying degrees to activate viewer memory and emotion. Authors design gaps, open areas and ambiguity to clear the way for audience contribution, increasing immersion into the world on the page. In Julio Cortazar's short story *House Taken Over (1946),* a brother and sister are driven from their home by an intrusive force gradually taking over room by room. Cortazar withholds details about who the intruders are and why they are there, creating a surreal and nameless antagonist, and this use of ambiguity enhances the symbolic strength of the work in enabling readers to imagine meaning. Andrei Tarkovsky cites the mutual dependency of the image and audience in literature, observing, "Prose too, of course, relies on the reader's emotion, spiritual, and intellectual experience as does all art. Not even the most naturalistic and detailed passages of prose remain within the writer's control: whatever happens the reader will perceive them subjectively."[26] This is not to say the author relinquishes control of content to the viewer, rather that artists methodically stimulate viewer memory and emotion to increase their ownership of the experience. In Alain Robbe-Grillet's novel *Jealousy (1957)*, the reader is personified as an elusive and unidentified narrator making observations and recalling events regarding his wife, a neighbor, and the house he occupies. The gaze of the narrator, and his repeated and measured descriptions of the space and characters, guide the audience experience, and subjective details about his mental state or objectives are hauntingly omitted when a crime is suggested. Analogous

[26] Tarkovsky, 176

to the extended duration in a cinematic shot, Jealousy uses repetition of prose to aid the formation of thorough mental schema in the reader. In addition, Robbe-Grillet provides a map in the beginning of the book (Figure 14) to augment the spatial dimension, allowing the reader to calculate and enact the events in the text.

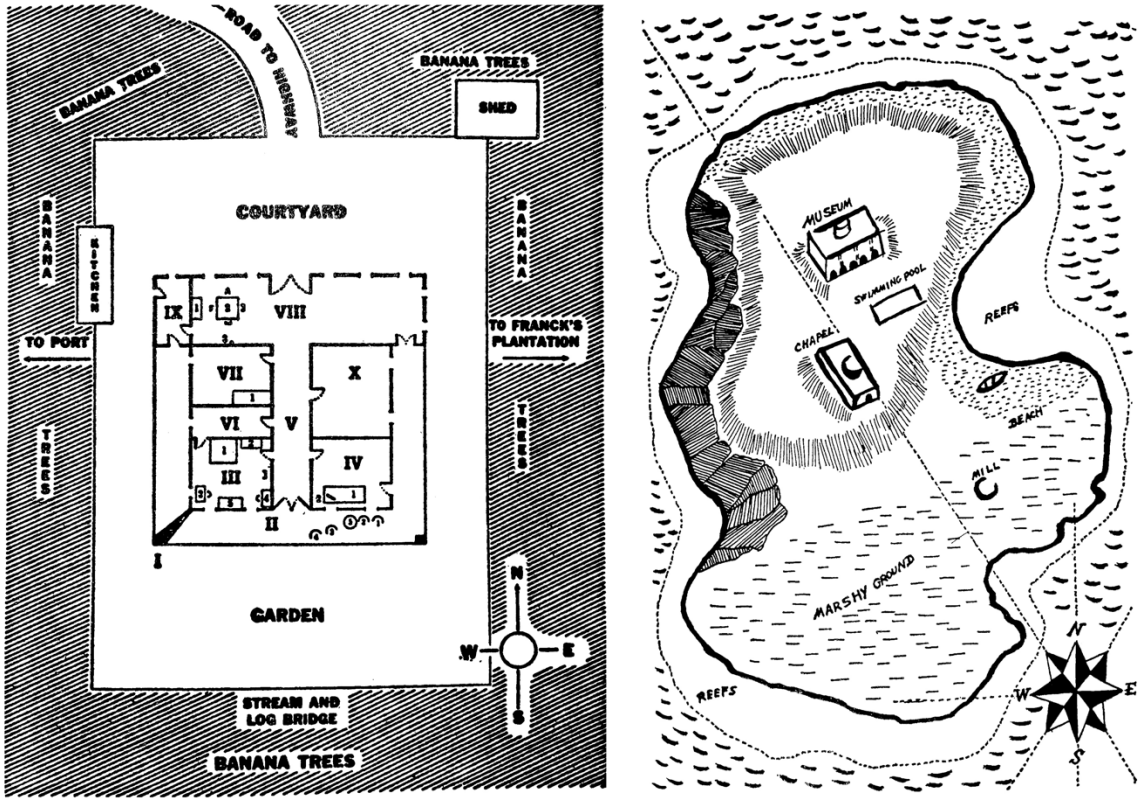

Figure 14: Map provided in Jealousy (left), and The Invention of Morel (right)

Similarly, Adolfo Bioy Casares' novel *The Invention of Morel (1940)* broadens the text by using a map and visual illustrations (Figure 14). The story follows an exiled man on a mysterious island where its inhabitants and spaces are gradually revealed to him.
He becomes infatuated with a character and realizes that she and the other people he observes are recorded images from an unknown past. Similar to Jealousy, repetition of

the scene depictions increases reader immersion through familiarity. Agency-environment in Casares' story is implied as we observe, from his perspective, the character negotiating the haptic, holographic environment mapped onto his reality and the impenetrable emotional object of his desires.

In defining a literary immersion by its variability and manipulation of perspective, Vladimir Nabokov's novel *Pale Fire (1962)* extends this definition in presenting a navigable fictional world requiring readers to browse passages in different areas of the book, each representing various points in space and time. The world of Pale Fire consists of a lengthy poem, and index, and extensive notes expounding upon the poem's creation and meaning through personal and historical stories. The characters and scenes are realized in detail, so audience inference stems from the relationship between the times and spaces depicted, the puzzling connections facilitated by the non-linear form. As author and historian Michael Wood notes about the novel, "We wonder at first, what kind of text this is; but soon wonder more seriously whose text it is, and what is the matter with him."[27] Nabokov offers a narrative architecture requiring readers to make choices and connections that ultimately form a dynamic image of the world.

The Final Image project mediates artistic immersion in the ways discussed thus far ranging from a metaphysical or pictorial immersion, to a literal and continuous image environment with viewer navigation controls. The VR display platform implemented in the work severs the viewer from their visible self, isolating their hand (controller) and

[27] Wood, Michael. *The Magician's Doubts: Nabokov and the Risks of Fiction*. Princeton, N.J.: Princeton University Press, 1995.

head (head-mounted display) as kinetic means and sensed understanding of the world. This individual-oriented characteristic of the apparatus originates from its inception as the stereoscope, a device positioned near or onto the face to display depth-images (Figure 15).

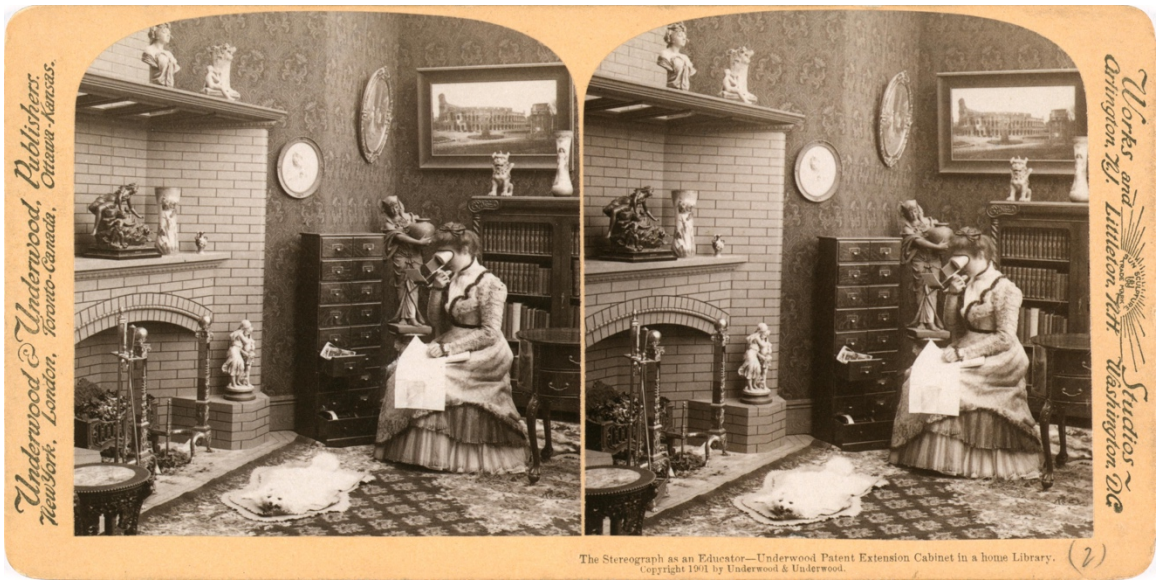

Figure 15: Stereograph card for display in a stereoscope

In opposing sound, widescreen formats, and stereoscopy in cinema in 1933, Rudolf Arnheim observes, "For a single spectator it would be easy to make a stereoscopic film. It would only mean taking two simultaneous shots of the same incident a couple of inches apart and then showing one of them to each eye. For display to a larger number of spectators, however, the problem of stereoscopic film has not yet been solved satisfactorily."[28] Although stereoscopic display in theaters has since been solved, the

[28] Arnheim, 11

wearable or embodied nature of VR display rehashes the seclusive problems of the stereoscope. On the other hand, the VR apparatus is not obligated to conform to a shared public exhibition convention, and instead, converges with the solitude of literature. In fact, a multi-sensory and embodied poetry is a reflection of a total phenomenological experience. Sergei Eisenstein prophesizes in 1944:

> Stereoscopic and color film are being realized before our eyes. And the moment is drawing near when, not only through the method of montage, but also through the synthesis of *idea, the drama of acting man, the screen picture, sound, three-dimensions and color*, that the same great law of *unity and diversity* – lying at the base of our thinking, at the base of our philosophy, and to an equal degree penetrating the montage method from its tiniest link to the fullness of montage imagery in the film as a whole – passes into *a unity of the whole screen image*.[29]

[29] Eisenstein, Sergei. *Film Form: [and] the Film Sense; Two Complete and Unabridged Works*. Vol. MG 10; MG 10.;. New York: Meridian Books, 1957, 254.

Section 3: The Final Image

The great function of the artistic image is to be a kind of detector of infinity…towards which our reason and our feelings go soaring, with joyful, thrilling haste.

Andrei Tarkovsky, Sculpting in Time

3.1: Research Objectives

Section 2: Project Framework conceptualizes several artistic problems through examples of creative work and philosophical inquiry. They can be summarized:

- Employ agency-environment relationship to enable immersion in an artistic image
- Negotiate authorship and composition in an interactive structure
- Reimagine the cinematic frame for an immersive media context

These objectives are central to *The Final Image* production and became evident through discoveries in preceding projects. Further, they intersect and correlate and are not addressed in isolation. For instance, viewer interaction is a component of viewer agency, which in part defines the immersive potential of an experience. In addition, concerns with agency/subjectivity, interaction, and frame are adapted to the specifics of VR display, and new questions of content and structure emerge. Several solutions are discussed.

First, an early problem was approaching the built-in subjective nature of the VR apparatus: how to establish a narrative vantage point with appropriate fictional justification since the audience is essentially wearing the camera. With whose eyes does

the viewer see the world, and how can their point-of-view be personified? The viewer is guided through a meditation in the opening, establishing a therapeutic context and framing the interactive scenes as visions conjured by cognitive exercise. In addition, a living body is imposed on the viewer with close-up sounds of controlled breathing foregrounding their visual experience. The combination of guided meditation and pervasive breathing form a *Rückenfigur,* an ambiguous 'figure from behind' through which viewers inhabit the spaces. In the photographic scenes, breathing stops, narrations are heard, and audience observation is guided through a character moving in the panoramic image. In this way, the *Rückenfigur* subsides, giving way to a more objective viewer position and liberating their perspective. By working in a spectrum between the first and third person, the identity imposed on the viewer is elusive.

Coincidentally, this connects to the second research objective in *The Final Image*: to negotiate interaction in relation to authorial and compositional coherence. Which parts of the work can the audience control? Direct interaction and displacement are eliminated in the photographic vignettes, and the duration of the scenes and audience position is fixed. This rigidity ensures that the entire lyric and rhythm of the voice-overs are heard, and in addition, metaphors the ceaselessness of recollection. In contrast, the kinetic and navigable characteristic of the three-dimensional environments reflects an active process of regaining memory. These opposing structures oscillate throughout the work, with the more passive, observable vignettes providing relief to the interactable 'game-space' interludes. Further, events such as animations and sounds are triggered in a specified

order. This structure maintains an authorship of montage and at the same time, permits the viewer to control the duration and viewing position of events.

The third research problem is adapting the cinematic frame (or shot) for a continuous and navigable visual field where there is no longer a divisive pictorial border, but instead, zones of intention or visual / kinetic limitation. First, the primary strategy for guiding attention in *The Final Image* is light: sources of illumination such as lamps, streetlights, and traffic lights are activated consecutively to direct attention towards tasks. In addition, audience look directions are predictable, so active lights are positioned in their periphery to ease transitions to the next focal area. As a result, visibility of the environment and depth of field increases temporally, and so does the audience sense of scale and immersion in relation to their actions is in the space. In this way, limits in how much the viewer can see, hear, and do, give way to a consistent compositional shape with variations from viewer choice.

The project objectives and solutions are oriented towards structure so artistic content must be addressed. *Section 2: Project Framework* problematizes research areas through several works of art, and those same works inspire the thematic and emotional materials of *The Final Image*. The spiritual quest in *Nostalghia*, the displacement of home in *House Taken Over*, the moral dilemma of *The Man from London*, the search for identity in *Dear Esther*; these themes and more inform the production. Similarly, my recovery from addiction and the moments before and after my transformation are an ongoing source of narrative. The section following this entitled, *Preceding Works*, discusses

subject matter in more detail and illustrates a trajectory as projects become more personal and complex with increased integration of concept and form. The question of content for all project work: transforming personal experience into the poetic materials of art.

The locations for the photographic scenes in *The Final Image* are both intimate and archetypal. The abandoned and decaying home represents the distortions produced by nostalgia, and the longing for a former self after transformation. The home is a well of memory and tragically transient, a manifest yearning for permanence. The movie theater location represents a zone of desire, a place of fantasy and recorded histories both fictional and real. The emotional draw is directed away from the screen and into the aisles and seats where the act of looking and personal memory takes place. Location selection was crucial for the piece and took place over several months. The house was chosen for its combined sadness and danger, as if a mysterious force had ripped it apart, leaving remnants of a domicile. The theater selected, where I saw screenings of rare movie prints and art films over the years, is small, dark and sentimental. The interactive scenes contain primitive representations of spaces, a cave, a domestic room, and a theater façade, all scattered like puzzle pieces waiting to be assembled, analogous to my process of self-examination and reform. An interpretive walk-though of the work:

> The work starts in total darkness, a voice guiding a meditation. Hundreds of bulbs fade in, faintly outlined, barely visible and strewn across the ground surface, suggesting the depth and scale of the space. A single light source turns on, illuminating the first task, then a second task, and third. Moving the rocks into place, one by one. With each stone a blurred and broken voice is heard,

becoming clearer as a cave forms and a portal appears as the final piece is placed. Moving through, dissolve to a field, a fence, a road, and an abandoned house. A man appears, recalling a memory of home, and searching the space, getting further away. Into the road. Fade back to darkness and return to the tasks. Reassembling the pieces of a domestic space: a couch, a chair, a window, a lamp, a T.V, the sounds of fragmented music and the voice becomes clearer. Another portal, to a corroded domestic interior and another monologue. He appears again, inspecting the deserted space, then dissolve into the dark space again. The assembled domestic space and cave quietly sit, and another task appears. Piece by piece, the assembly of a movie theater, a marquee is erected, the bulbs lift and spin in a cloud then land on the façade. The portal appears. In a small cinema theater, the man sits, staring into the dark from the seats. No screen. Prose is recited again, the parting words, the movie screen appears, and he reaches out. White dissolve to the space where the cave, the living room, and the theater are now shattered black forms, floating in the white space. A negative of the scenes before. Back to black.

Additional discussion regarding the primary research questions and content appear in this document under the production details of *The Final Image* and the works leading up to it. Research moving forward will focus on the overarching problems of fusing content, architecture and mechanics. One area that needs expounding is the problem of perspective: to personify the viewer with more variation and detail, and to futher disrupt / embrace the subjective nature of VR. In addition, future research will experiment in

integrating elastic and fixed states more closely to solve the segmentation of these structures in *The Final Image*. Finally, the interactive spaces should be deeper in volume and narrative to enable and reward more exploration from the viewers. Ultimately, production of *The Final Image* artwork established a poetic and technical foundation for the years of research ahead.

Figure 16: Various moments from The Final Image.

3.2: Preceding Works

I was drawn to stereoscopic imaging by a desire to bring new visceral dimension into my work - to fuse the image more closely with the body of human vision. In addition, the pursuit was inspired from proximity to colleagues working in the medium, and readings in seminal phenomenological texts from Deleuze and Merleau-Ponty. Following a year of experiments and collaborations to gain expertise in stereo-3D image capture and display, the first project, *A Mountain in Two Eyes (2012)* began with writings on vision and depth and became a stereoscopic-3D video and sound work about the presence of Mount Rainier, an iconic peak in the northwestern US, through youth and later trials in adulthood. An excerpt from my text written for the piece:

> *Depth is a sensed volume, an agreement created through difference. A synthesis of vision streams, each bearing a separate image to be delivered and assembled. It is a simultaneous difference-agreement, the resolution of two images, guided by intention. A choice, a focus, a region of interest pulls our vision, and our eyes converge. My body is positioned on the terra, a seeing self, seen from self, seeing. That which confronts me, defines me. My scale appears through the volumes I sense. These volumes. These depths. Sensed through two eyes oriented horizontally in a skull, bending on a column of vertebrae, blooming a network of nerves, and creating an image of sense. My limbs, skin, eyes, ears, nose are gatherers, painting their sensuous findings in the mind's gallery. When confronted with a thing, these receptors open up like sails trying to catch wind. If the thing is small, like a coconut, I feel it from all sides, holding its volume. I smell its unique odor hovering closely around its form, and hear its throaty timbre*

sending small waves of sound when tapping it with fingers. These interactions define our scale in relation to the thing – the coconut scales us. We see it. Twice. We see it twice, from two sides, from two eyes – at once. Gripping the hard, round shell, we interact, but with two eyes, we enter its volume, its depth, its life.

Viewers use glasses to immerse in 3D landscapes displayed on a large surface where voice recordings were also heard (Figure 17).

Figure 17: Exhibition view of A Mountain in Two Eyes.

Figure 17: Terrain map of Mt. Rainier with camera positions.

Stereo-pair images were captured on location using deep interaxial (Figure 18), placing the mountain and its surrounding range into stereo-depth for the first time and creating a visceral experience of a landscape typically limited to monocular depth perception. In addition, voice passages about stereo-depth, childhood and transformation were recorded during shoots on the mountain, and this personal prose acts as an observer and point-of-view for the landscapes, an aural lens that colors the images with intimate reflections. Although the work provides a subjective vessel to see the landscapes and a novel sense of depth, the projections themselves do not contain an intimate or spiritual quality and remain vacant in many ways.

The work that followed, *Between This Image (2013),* embedded stereoscopic-3D video and sound into constructed scenery in the exhibition space, inviting viewers to navigate the physical and projected content (Figure 19). The piece depicts a video-recorded figure facing the impossibility of becoming conscious or self-aware and was based partly on Magritte's painting, *La Reproduction Interdite (1937).* In addition, text content informed by Adolfo Casares' novel, *The Invention of Morel*, was written then recorded by a voice actor. Clips of the text were edited and categorized by parts of speech then sentences algorithmically assembled. *Between This Image* increased the poetic and immersive complexity for the audiences, advancing the treatment of text materials and fabrication. Stereoscopic-3D images are projected onto a transparent scrim surface to create a holographic effect between the real space and the media. The two-way mirror in the rear of the set piece reflects the scrim and displays images from the monitor mounted behind, creating a recursive reflection moment.

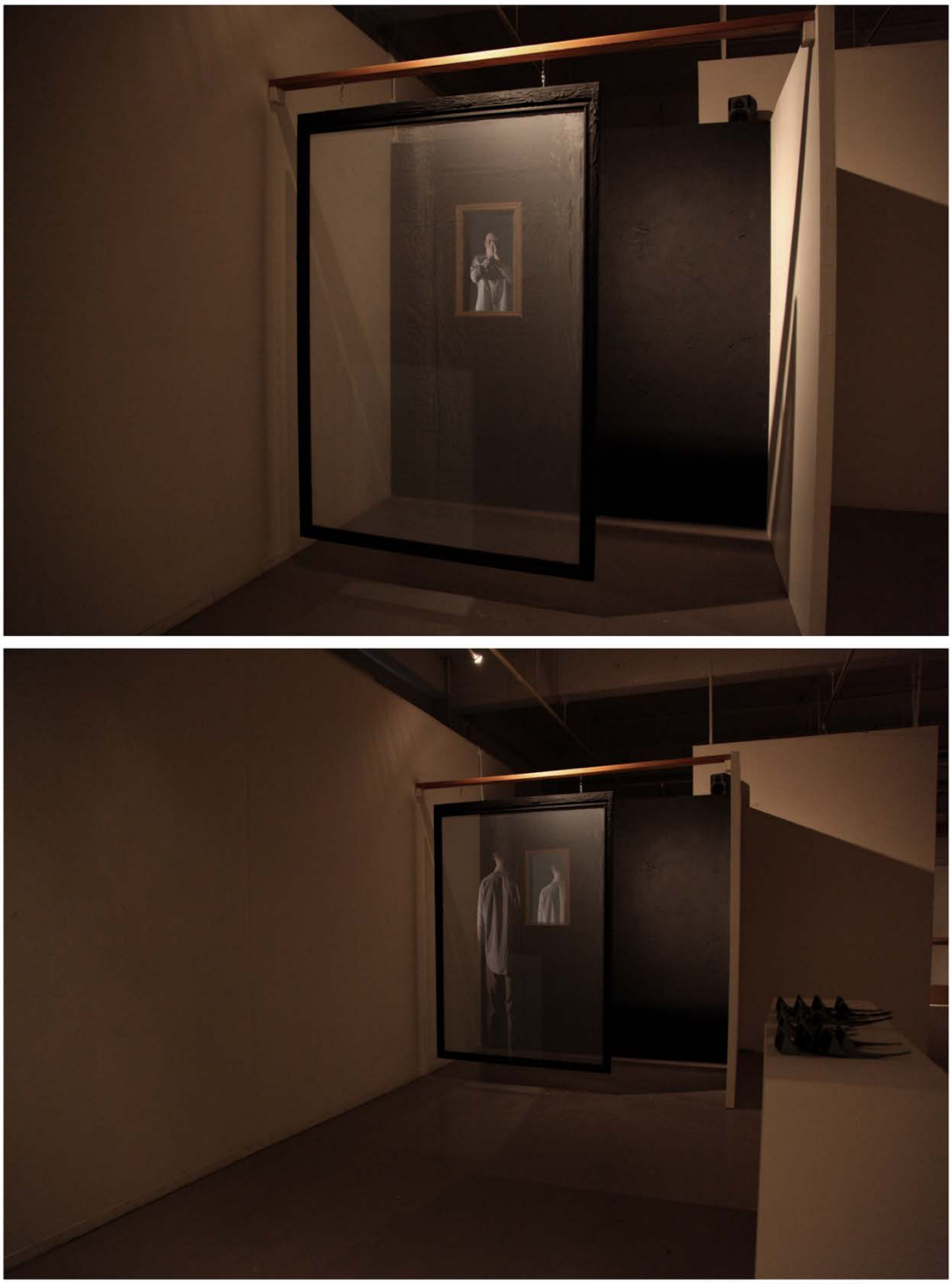

Figure 18: Exhibition view of Between This Imag.

Similarly, the artwork *I'll Make Myself a Memory (2015)* is an immersive moving-image and sound theater using large-scale projections on fabricated walls (Figure 20). The physical position of a viewer and their look-direction determines the cinematic experience at any given moment since all three screens can never be seen simultaneously. The piece was conceived first through writings about the temporal warping and isolation experienced during an extended bout with addiction, and the title is taken from the Samuel Beckett work, *The Unnamable (1953)* where a series of run-on and broken sentences express a riddle monologue of mental confusion. A windy window, a mirror and sink, and a bed are a triplet of object spaces depicted in the three video channels and they detach and become visually continuous at different times in the work.

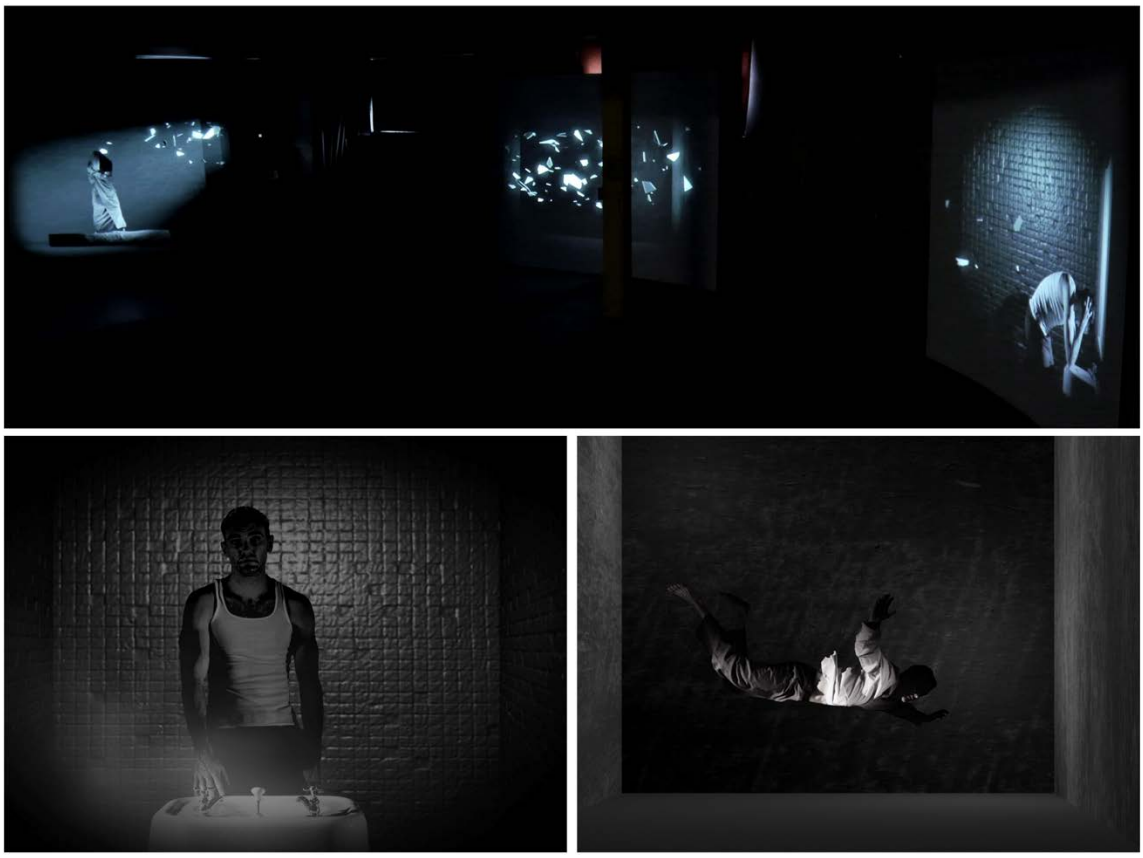

Figure 19: Exhibition view (top), video stills (bottom).

The subsequent work, *Memoir ll (2016),* is an autobiographic cinematic work developed for virtual reality. Expanding on previous project techniques, the work uses personal experience as the narrative lattice and the technology provides a continuous and navigable visual-sound experience for the audience. The project commenced by gathering my father's hand-written journal entries taken during my hospitalization in late-stage alcoholism. The entries were typed and edited, taking poetic liberty to broaden and compose the text for viewers. The content was read and recorded by a professional actor and used as voice-over for the visual experience of the piece. An example passage:

Adapted Entries

Got to the room, now sitting with you.
Downtown. The bigger building.
With you.
I'm exhausted, slept a little last night.
Cancelled my appointment to come today.
You're asleep.
You're so sick.
Looking better, and comfortable.
A towel over your eyes, food nearby.
She put cold water on the towel.
Woke up for a minute. Shifted a little.
The room is hot.
She lowered the temp. You told her to lower it more.
She does. Selfish.
You're not eating much. I'm hungry.
My head hurts, lack of sleep. Will go home soon.
You are so selfish right now. The way you ask for stuff.
But then you say thank you.
She says it could take a few days before it all wears off.
You are taking new medication.
Your throat moves, big gulps, dribble. Thirsty.
I told you about the results, about your condition.
You tell me not to talk about this now.
It's too disturbing.
I stop talking.

Original Entries

I'm sitting in the hospital room
410 in Harborview
With Martin.
I'm pretty exhausted. Slept a little last night.
Got up and going
Cancelled my apt in Redmond this morning.
Martin with liver failure
He looks better & comfortable
Has a towel over his eyes
Breakfast tray next to his bed
Got the aid to put some cold water on the towel over his eyes.
Just woke up for a minute
Hot hospital room, temp set up to 80
Aide put it down to 70. Martin asked it to be put down to 60. Aide did this. [Selfish]
Martin filling out the menu for today and tomorrow. Not eating much.
[Treating the family]
My head is hurting. Lack of sleep.
Will go home soon.
Martin didn't sleep well last night
"was in pain"
(Martin seems very selfish right now -)
The way he asks for stuff
But then he says thank you.
Just checked with nurse, Deborah
Said that it could take a few days before he went into withdrawals from the alcohol
[acites draining fluid]
He's getting some diuretic medication
To take the fluid off
I told Martin about the ultrasound
About the fatty liver
He looks and said that my talking about this with him was disturbing

The primary narrative space in the work is a hospital wing where viewers move through the halls and rooms accompanied by the voice over. The 3D models for these scenes are based on recollections of the space during the hospitalization, and details such as scale and architectural complexity were modified for dramatic effect (Figure 21).

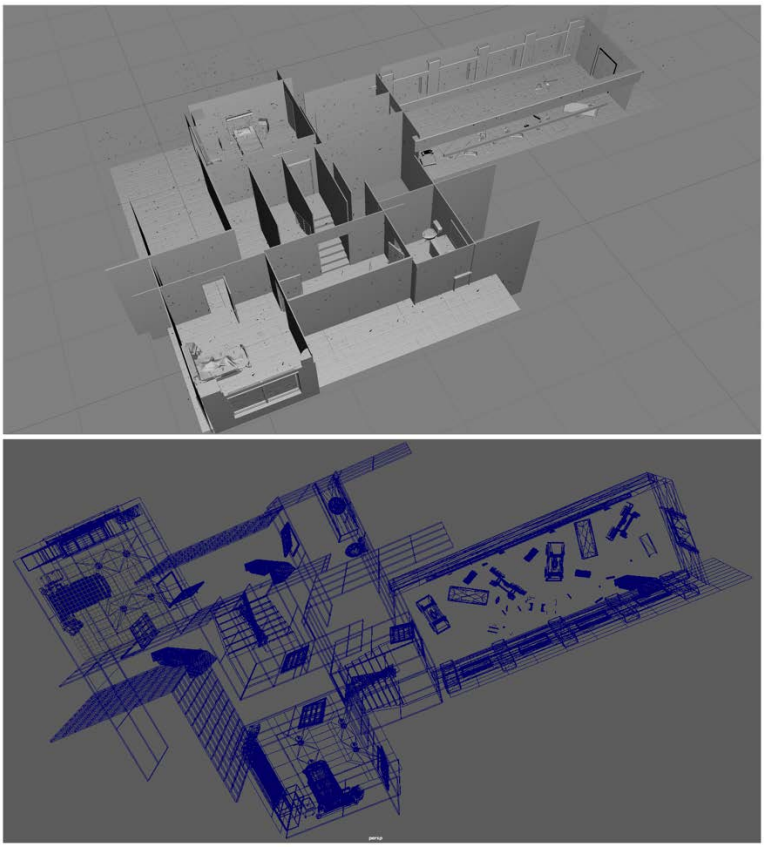

Figure 20: Hosptial model in shaded and wireframe view.

For example, doorways and halls are narrowed to increase the sense of confinement, and pathways through the wing bend to decrease the audience ability to anticipate the spaces in front and behind them. To further augment the claustrophobia of the hospital wing, a spot light effect is applied to limit the visibility of the space (Figure 22).

Figure 21: Screen shots from game: Hospital from viewer perspective (left), corridor scene (right)

The voice-over recites the journal entries and a quiet layer of city sound is added for depth and context. The work periodically transitions to dark ambiguous hallways lined with doors, some of them slivered open to reveal patients in examination rooms. Voice recordings from different actors were filtered for distance and occlusion then layered into the scene to give an impression of other patients in the vicinity. These sudden interludes represent the blackouts, confusion, and hallucinatory episodes experienced during detoxification.

One problem that this work posed was how cinematic techniques could be applied in the VR platform in a way that both expanded conventions and embraced the potential of the new display technology. For example, to create an analogy with the film frame in cinema, a travel path was created which viewers could follow as if moving along a rope through the space (Figure 23). This allowed continued freedom in head and body movement and at the same time, tethered the viewer to a specific line of travel, or frame. After designing and modeling the visual architecture for the work, the travel paths were created and applied in the scenes, each one framing different areas of the space at various times.

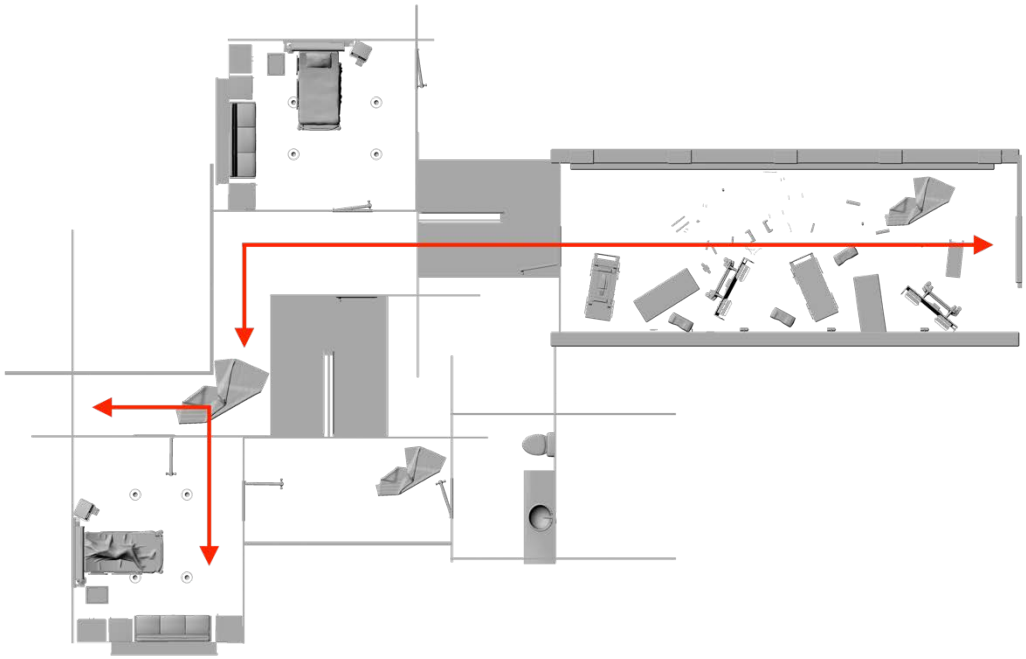

Figure 22: Top view of hosptial wing model with animation paths (red)

In addition to the regulated viewer movement, visual breadcrumbs and lighting were used to guide and compel audience movement in the spaces. In the hospital wing, selected doors are left open to invite the viewer in and bedding is strewn on the hallway floor in the direction to follow, and in the interlude scenes, images of a gowned patient are posed at different points in the corridor signaling where to look and travel. The final segment in the piece provides relief from the limited visibility and tight architecture as the viewer enters a large grassy field shot in spherical stereo-3D.

3.3: Writing and Voice Recording

The work's composition is ordered in an emotional logic similar to music or poetry where pieces of story are depicted in voice-overs and visual objects distributed throughout the scenes. The work opens with a guided meditation to contextualize the experience as a reflective exercise, placing the viewer in literary second-person perspective:

> *Close your eyes and breathe deeply. Let the air clear your mind. The breath expands the space, sweeps the debris, and settles your thoughts. Observe your mind working. Let it create. It does this without you. Within you.*

The primary narrations in the piece are a series of verses in first-person perspective, passages about returning home, the recovery of memory, and confronting change. These sections are a voice-over component for the immersive photographic scenes in which audiences observe the protagonist exploring the spaces.

> *I come back. And everything's gone.*
> *The window. My window. From when I was young.*
> *I look in, but nothing.*
> *And out. I look out. Lost.*
> *I am looking at the present. The things in front of me right now.*
> *The shapes, depths, colors. As I stare at them, an inner vision grows.*
> *A separate set of images flickering behind my eyes.*
> *An unseen sight, nestled in the volumes inside of me.*

The verse in the second image mentions forces and circumstances that have displaced the narrator from his home, and the specifics are elusive. In this way, ambiguity in the underlying language of the work enables audiences to process the content with their own intellect and memory. The narration continues:

Rubble. Ruin.

Wood and dust and metal. Concrete.

Cracked slabs and boards leaning shredded. Stacked.

Dirt. Weeds, puddles.

Mounds lay wrapped and lumped.

My home devoured.

A nightmare where everyone, everything turns its back.

Flips inside out. Familiar just enough to let me in.

They crack the door to slivers of memory, blazing and fast.

It's in me. My home.

It flows with my blood and lives in my muscles.

Is this the same air? Same mind? Same memory?

And now. Others are here now.

I should've stayed. I couldn't have stayed.

In the final passage from the work, the narration crescendos:

Another passageway. The final image.

Not as I thought.

It fires away, relentless flickering stillness.

Home.

Not the one I expected. And peace is coming.

A slow wash in speedless movement.

Arms releasing. Opening.

Legs releasing. Opening.

And mind releasing. Another passage, the first image, back to the beginning.

Those dark spaces. The gaps between.

The openings divide in the wreckage. Passageways. Passageway.

A doorway forms without me.

Within me. Without me. Within me…

The text was read and recorded at a studio facility, and in addition, breathing sounds were performed and captured for use in the work (Figure 24).

Figure 23: Terence and Angela Kelley at Jack Straw recording studio, Seattle.

3.4: Sound Design

There are three categories of sound activity in the piece heard both separately and together at various times. First, are the voice-over narrations which are composed with layers of concrete and processed sound to provide ambiance. For instance, in the first photographic sphere we enter, an outdoor field-recording fades in with the sound of birds and breeze. This recording was captured on location during photography and minimal processing was applied in post-production to maintain a natural and diegetic sense (Figure 25). This ambient layer transitions to another outdoor recording with more pronounced wind from a different location in order to subtly increase the sonic intensity of the scene. The sound of a car passing on the nearby road was isolated in the recording, then treated with signal stretching and reverb for a surreal effect: an incidental and referent sound becomes a point of abstraction and departure from the familiar. The voice-over recording is kept conventionally dry, then mixed with the location sounds and processed car audio.

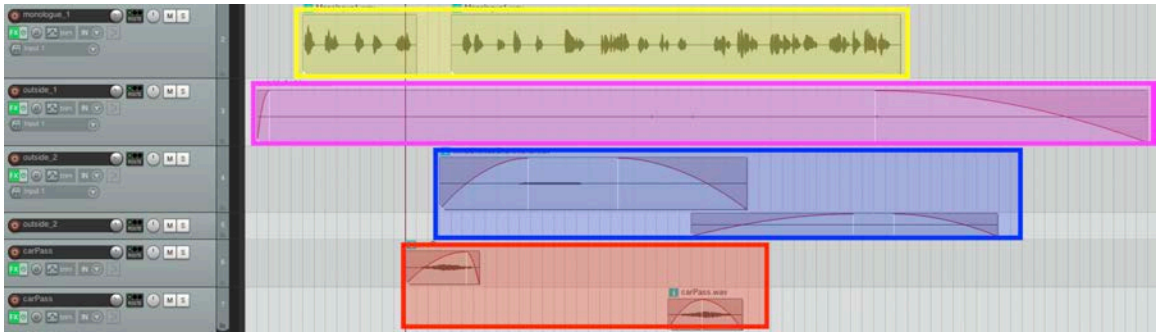

Figure 24: Mixer view of monologie tracks: voice over (yellow), outdoor ambient (pink, blue), car passing (red)

This process is relatively similar in the two other narration scenes: layering voice-recordings with multiple ambient sounds and effects to create a subtly surreal depth and immersion for the spherical image segments. Additional voice over is used without

ambient layers in the instructional scene and prologue section of the piece to create a more sparse and objective sound profile.

The second area of sound composition in *The Final Image* are the clips which are triggered with animations. Each interactive scene presents the viewer with the task of reconstructing a doorway with multiple objects, and each object has its own sound played back when triggered. To illustrate, audiences encounter building debris in the third interactive space and are tasked with moving the pieces into place to form a passage. Each object they move triggers a sound clip (Figure 26), and in addition, a sound clip is triggered each time an object completes its animation, so viewers clearly know when to move on to the next task.

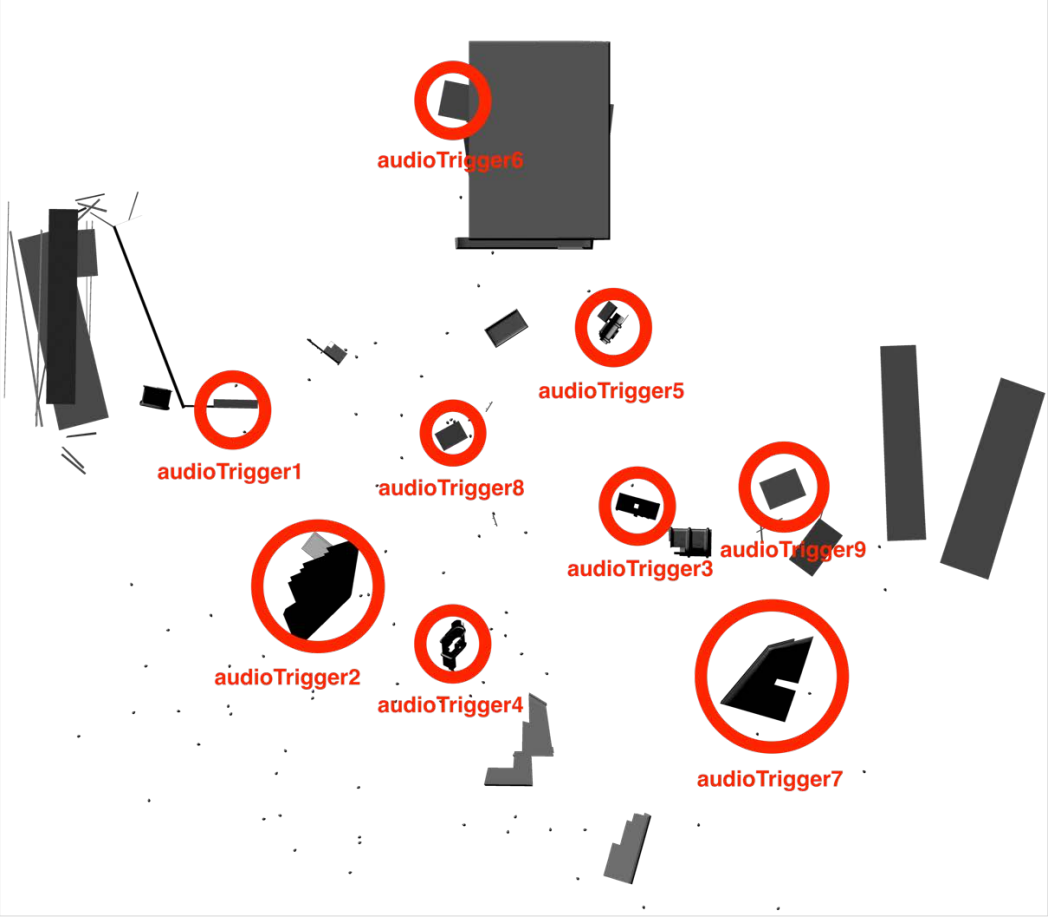

Figure 25: Top view of theater scene with marked audio trigger pieces.

The sounds are designed in relation to each other and to anticipate the length of object animations. For instance, if an animation is 15 seconds long and its triggered sound is 30 seconds long, it is likely that the sound will continue to play after the animation completes, creating an overlapping transition with any subsequent animations (Figure 27). Conversely, clip lengths are at times designed shorter than the animations for compositional variation and increased periods of rest. In addition, sounds are variations on a theme: each group of objects has a sonic character in which the varying individual clips adhere.

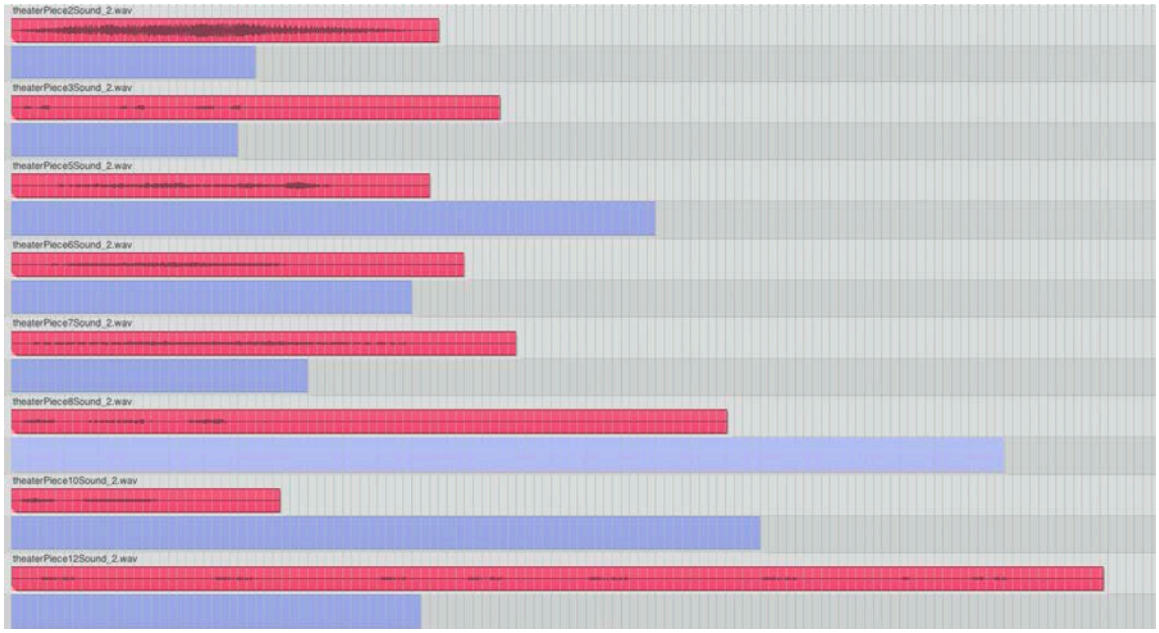

Figure 26: The lengths of sound clips (red) varies in relation to length of animation clips (blue)

This durational relationship between the animation and sound clips is applied throughout all interactive spaces in the piece: when viewers activate select objects in the scene, sounds of varying length and composition are triggered to create a dynamic soundscape. Each set of interactable objects has a sonic theme that ties them together and contains processed voice content sourced from the narrations. The voice is granulated to obscure

direct legibility and serves to forecast the monologue that follows it. For instance, as the viewer animates objects in the space, they hear obscured parts of word, stammers, and vocal sounds in the triggered clips. After completing a doorway, they enter the photographic sphere scene and hear the resolved legible version of the voice-over narration. This design conveys the struggle to recollect and assemble memories and identity after trauma.

The third sound component in *The Final Image* is the breathing and clock layer heard during the interactive interludes. The persistent breathing sounds refer to the audience inner state and rhythm that guides the haptic-visual experience. The ticking is a framing device, eluding to the abstract space where the guided meditation originates, perhaps a room with a clock. As sounds emerge interactively in the scenes, the clock is a constant presence to ground the viewer. Similarly, the breathing sounds are constant and impose embodiment onto viewer experience to accentuate the head-mounted display. The even and controlled breathing combined with the clock superimposes a polyrhythmic grid that reminds the viewer of their body and space during the experience.

The breathing was performed by an actor and recorded at close range with low gain to capture the strongest signal with least noise possible. Incidental mouth noises were removed, and high frequencies filtered for a helmet effect for the viewer. A single clock tick was recorded then algorithmically sequenced with two percent variation in pitch and at one second intervals plus or minus 0.05 seconds of random margin. These shifts act to offset uniformity in experience of pitch and time.

3.5: Location Photography

The three image sphere scenes in *The Final Image* appear as narrative vignettes in between each interactive segment and are created with stereo-3D location photography and several stages of post-processing. The high fidelity of each spherical image environment acts in counter-point to the geometric sensibility of the interactive scenes, providing relief and representing a moment of clarity, legibility, and resolution (Figure 28).

Figure 27: Spherical image mapped onto geometry.

Artist Paul Berger's use of continuous photography in his series, *Panorama (2008-10),* express infinite images within the finite of the frame. Referring back to Bill Viola's quote about experience, Berger's pictorial field liberates the confines of the camera view, capturing the entirety of the space, on all sides. Curiously, in one of his panoramas taken in a forest area, a dog is repeated in three positions as it moves across a creek (Figure 29).

Figure 28: Berger's panorama of forest scene with dog repeated three times.

This is possible due to the nature of panoramic photography where multiple single-frame images are merged into a span. This inspired a useful architecture for the panoramic still-image narrative in *The Final Image* where repeated instances of a character could appear throughout an image and the viewer could freely observe in any order.

First, houses were scouted and auditioned for qualities from the narrated text: lost and scattered memory, decay, nostalgia, natural light, and windows. Upon selection of the location, a large portion of miscellaneous litter was removed, leaving mostly decay debris and remnants made of natural materials. The choreography for the character depicts reflection and searching in several gestures distributed in the panorama. A general overhead diagram was created for the location shoot (Figure 30):

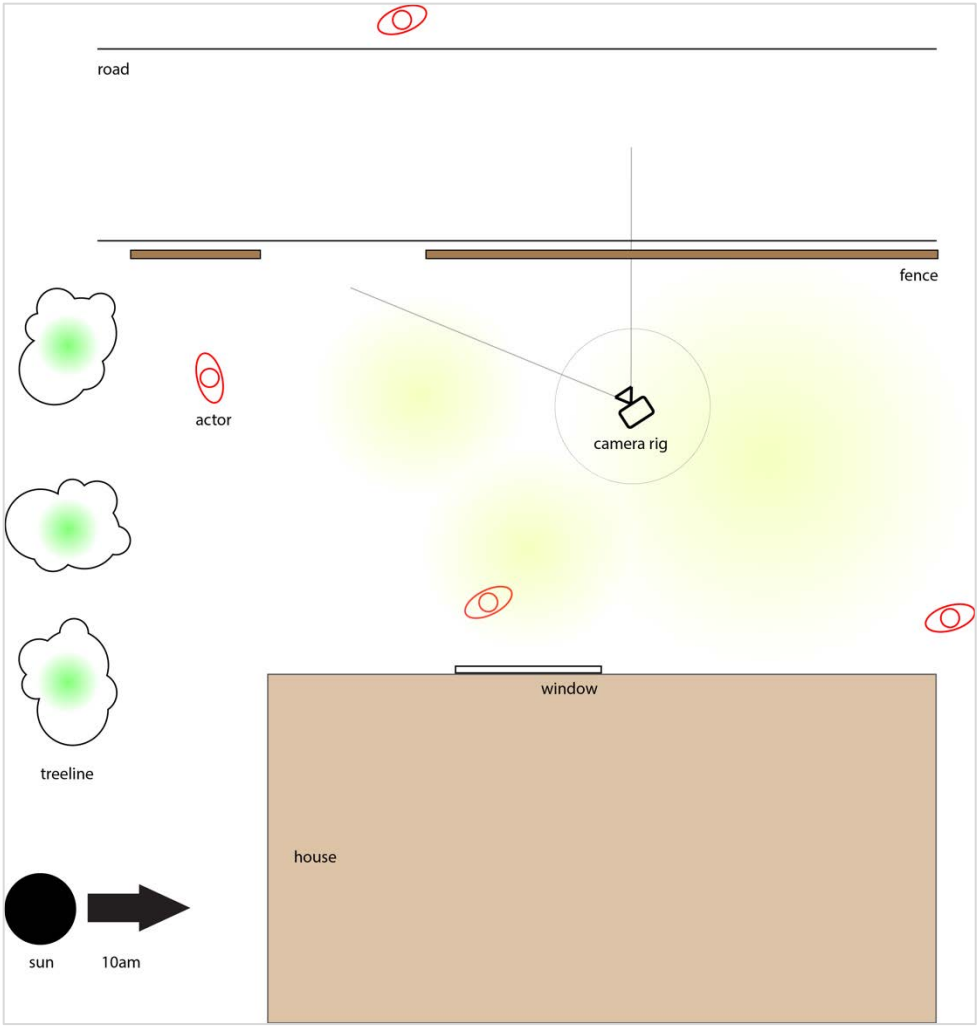

Figure 29: Overhead shooting plan for outdoor scene.

Two of the three spheres were captured on the same rural property from the inside and outside of an abandoned house, and the third in a movie theater. The capture technique for each location utilizes two cameras mounted in a beam-splitter casing attached atop a tripod. For each camera, images are taken in a full horizontal rotation with twenty to thirty percent overlap, then the sweep is repeated on the vertical axis with similar overlap (Figure 31).

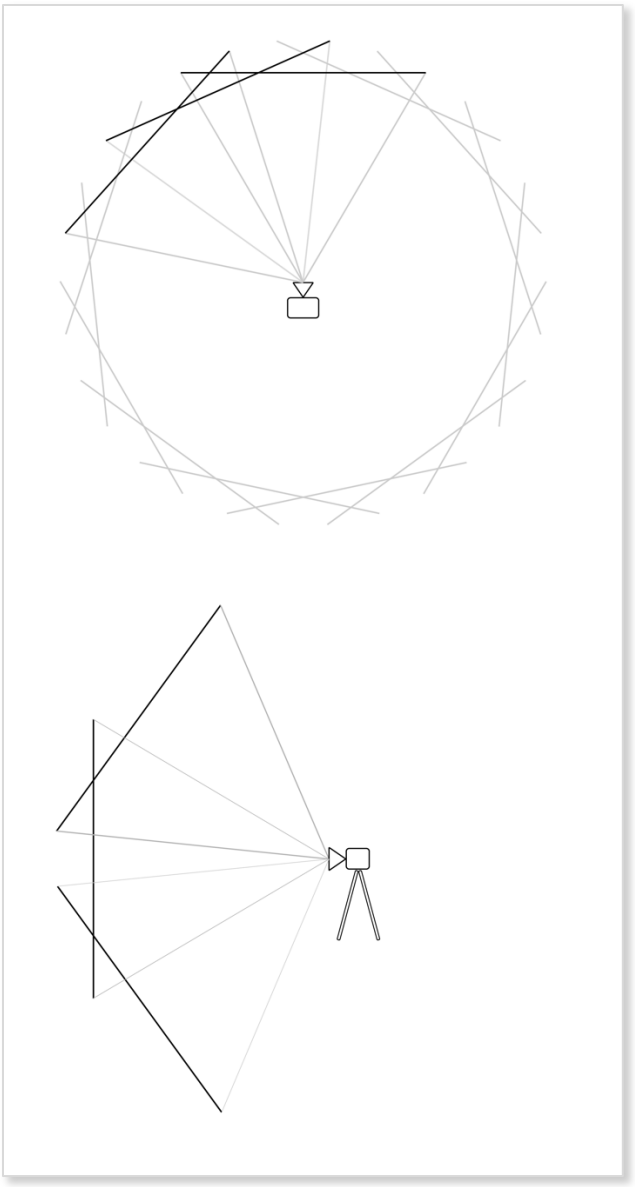

Figure 30: Horizontal sweep(top), vertical sweep (bottom)

The space was photographed without actors first to obtain an image foundation, then the character was directed and captured at various positions on the center horizon. For the interior of the abandoned house and the movie theater, the interaxial of the two cameras was lowered from the conventional interocular distance (0.62cm) to accommodate the small space and decrease the stereo-depth to increase viewing comfort.

Figure 31: Various production stills from location photography (photo credit: Adam Hogan)

The location photography yielded approximately two-hundred and eighty images for each of the three scenes (one-hundred and forty images per camera), a large yield due to an image-overlap of nearly thirty percent and an extra row of images on the vertical axis to ensure accurate results in the image stitching process. First, the content was batch processed to rename and correct inconsistencies in color and frame that can occur when using the camera housing. The image stitching software uses a control point algorithm to discern neighboring images and implement map warping (Figure 33).

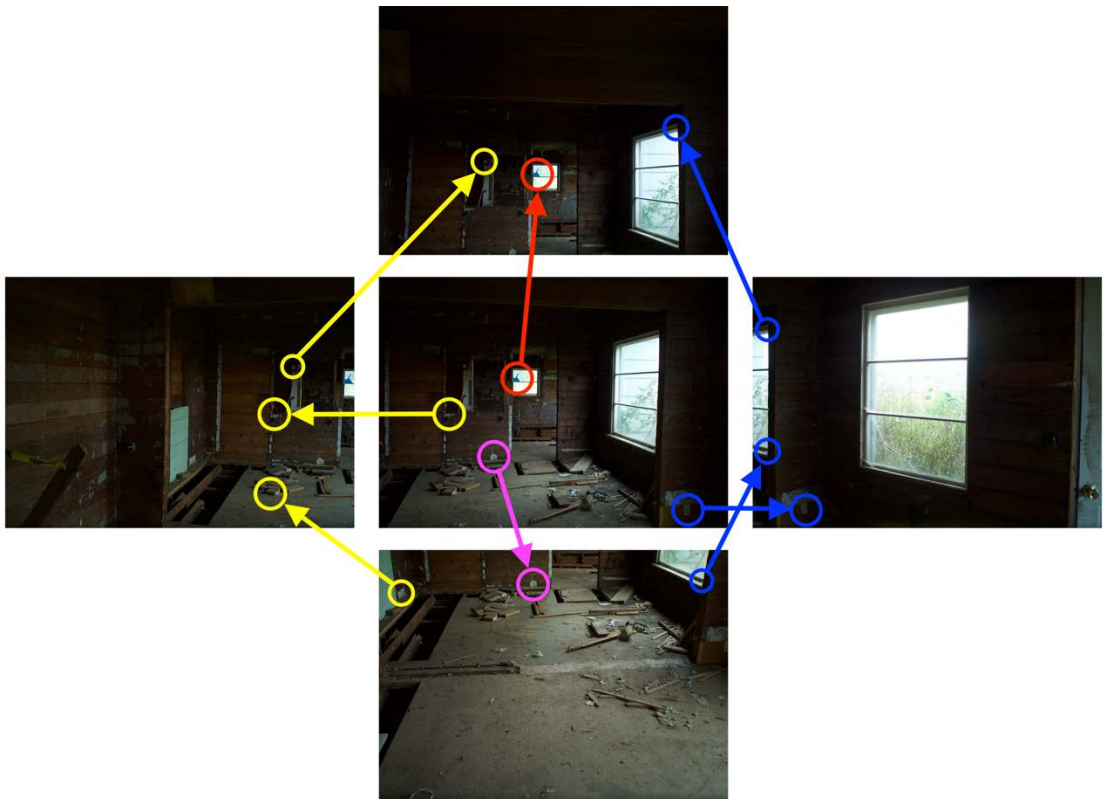

Figure 32: Sets of pixel-based control points determine adjacent images

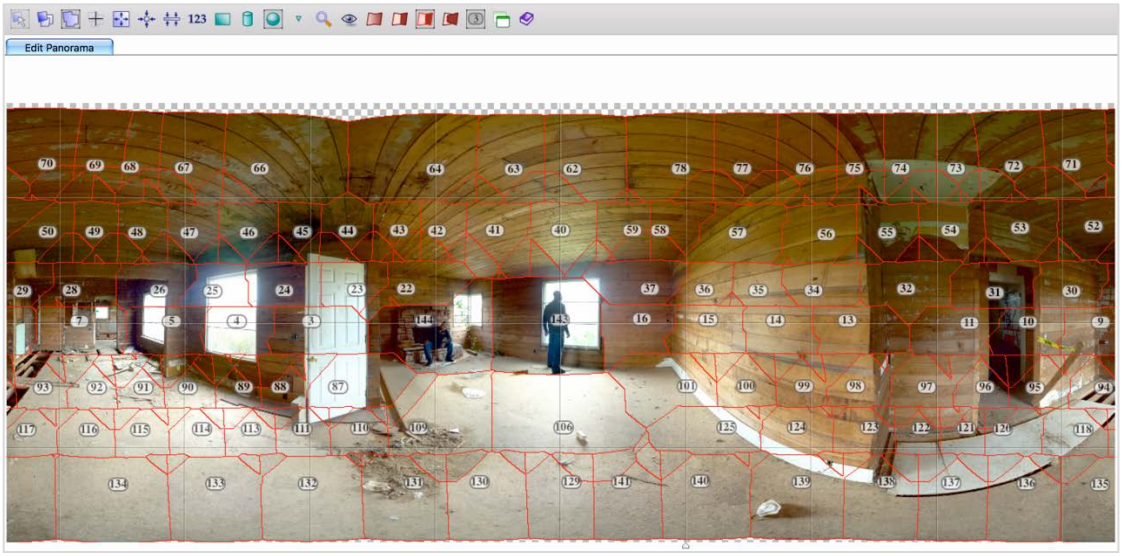

Figure 33: Exposure boost in software aids the pixel-finding / spherical alignment process. Tiles with character action are inserted into the set after a foundational image field is established

To increase the accuracy of the process, images are introduced incrementally by row from top to bottom or vice versa. This reduces the risk of the algorithm confusing information between images in disparate parts of the spherical field. In addition, a temporary exposure increase in the pictures aids the software in finding pixel-based control points. Image tiles containing the actor are added to the set and evaluated by the software for integration into the sphere (Figure 34).

Once the left and right camera image sets are entered and the software has completed the alignment and warping process, the raw results are monitored in a head-mounted stereoscopic display, such as a VR headset, to find erroneous stitching areas in the image field. Flaws such as misinterpreted placement of the images are fixed by manually assigning control points on neighboring images and repeating the software evaluation. Additionally, disruptive visible seams between images can often be resolved by applying masks to portions of the image containing discernable shapes with consistent color such as doors or windows (Figure 35). The masked images remain in position and are blended with neighboring images with a feather effect.

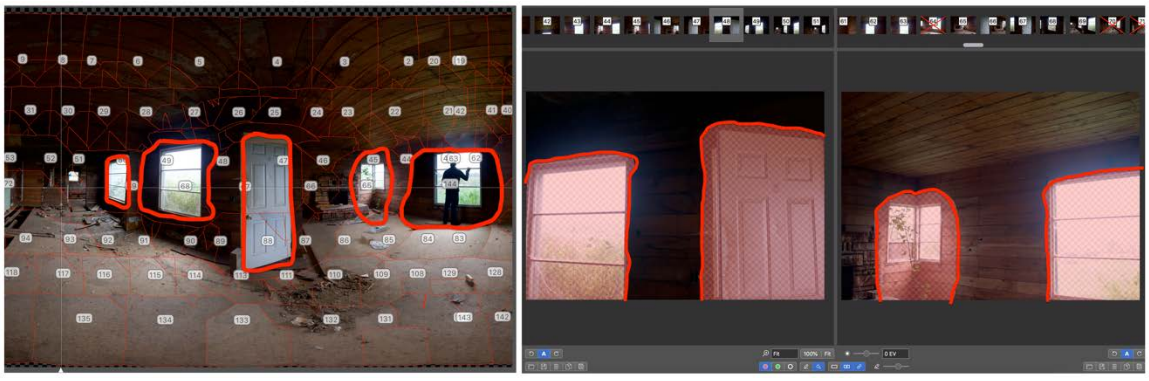

Figure 34: Masking areas of consistent shap and color to reduce visible seams in stitching.

Figure 35: raw stitched image (top), post-processed stitched image (bottom)

Spherical panoramas for the left and right camera are rendered then loaded into an image editor for color correction and compositing. For the interior scene (Figure 36), the color and illumination profile for the images is high-contrast and desaturated to draw attention to the windows, shapes, and obscure the dark areas for tension. As before, the content requires revision and quality tests in the head-mounted display where it will ultimately be shown. This is accomplished first by mapping the right and left images to separate geometric spheres with identical coordinates and positioning the viewer camera in the center. Culling masks are applied, blocking the right image sphere from the left eye and vice versa essentially using the head-mounted display as a stereoscope (Figure 37).

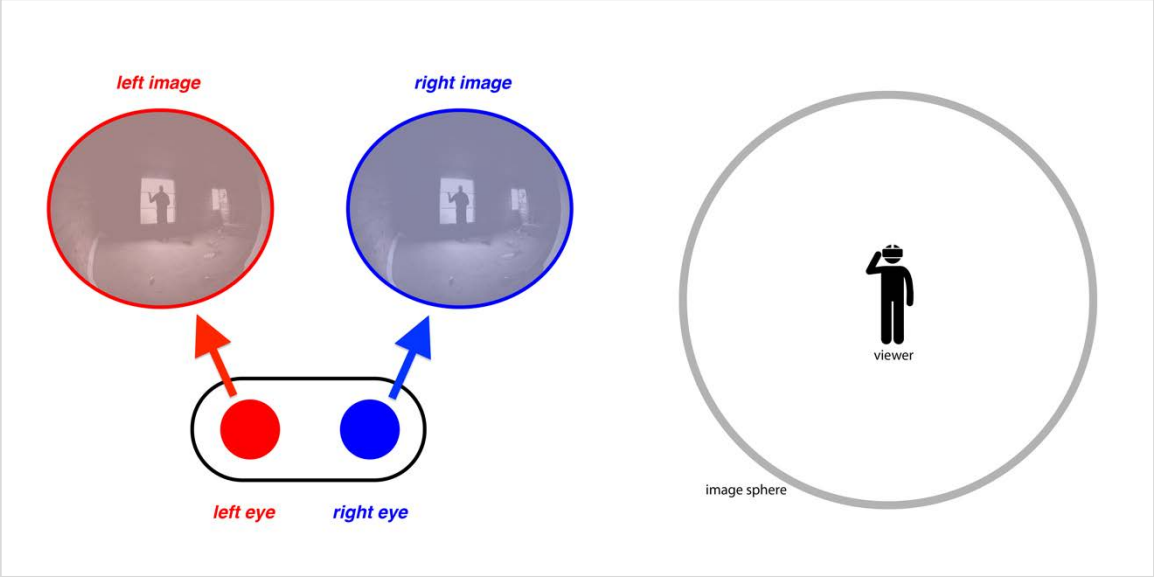

Figure 36: Image culling relicates a stereoscope for viewing stereo-3D images

Viewers enter the sphere space and are surrounded by a continuous stereoscopic-3D image field. The character is seen, frozen in still image gestures that change position around the viewer periodically similar to a slide show accompanied by narrations (Figure 38).

Figure 37: Panoramic image sphere illustrating position of character gestures seen over time

3.6: Modeling and Animation

Objects and buildings in the interactive scenes were designed and animated in a 3D modeling program then assembled, lit, and programmed in the game-development environment. The three main structures that viewers assemble are a boulder cave, a domestic living space, and a movie theater exterior.

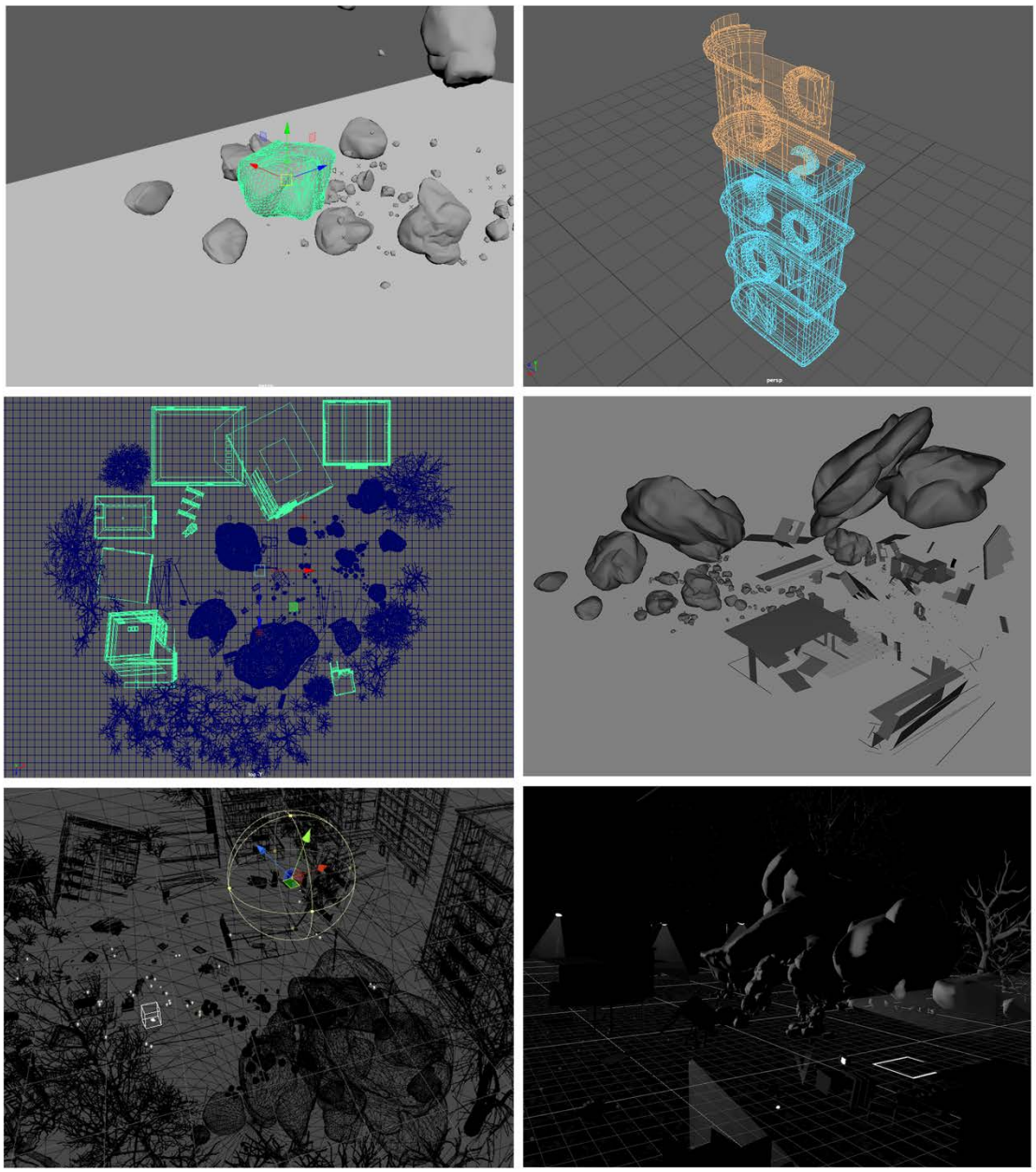

Figure 38: modeling and arranging geometry (top), whole scene (mid left), shatter animation (mid right), assembly and lighting in engine (bottom)

First, the geometry was modeled and constructed, then a shatter animation was created to disperse structural pieces as debris in the scene. The animations were reversed since viewers encounter the forms scattered before assembling the structures. The hand and wrist representing the viewer controller in the piece was modelled and animated, then light source objects such as work lights, streetlights, flashlights, security lights, and traffic lights were created. The keyframe clips for all moving geometric content were recorded and stored in the model files for access and manipulation in the game-development environment. The cave, domestic space, and movie theater pieces remain flexible for global scale adjustments, but the position of the objects must be maintained due to the cascade relationship of the pieces to each other: since the structural pieces lands on top of one another during the shatters, objects must adhere to a sequence when being animated in the game-space (Figure 40). Conversely, the light objects such as work lights, streetlights and so on, have flexibility in position to anticipate the highly iterative lighting process when designing the interaction order in the scenes.

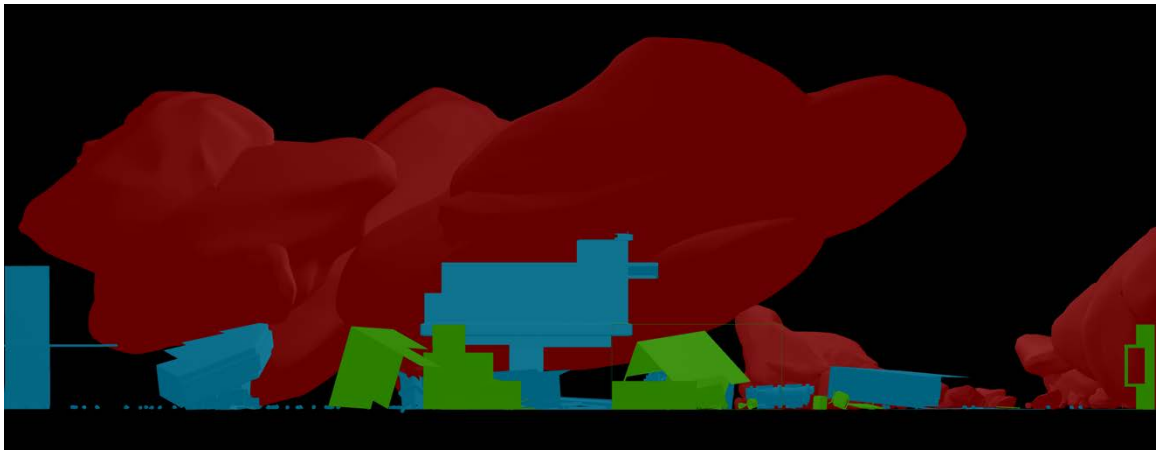

Figure 39: Debris stacked

3.7: System Overview

In technical terms, *The Final Image* is a time and interaction-based content rendering and playback system for virtual reality display. 3D models, lights, sound files, and photographic image materials are assembled in a game development environment, then activated in various ways using runtime code and physics components. Events are triggered on specified conditions, viewer input, or at given clock times, combining a fixed duration and dynamic structure.

To illustrate an event sequence, in the first interactive scene, the viewer is confronted with a boulder sitting static, illuminated by a streetlight. When the hand-held controller is pointed towards the lit boulder and the trigger is pulled, its animation component is activated, the associated sound is played, and the object moves along a keyframed path. If the controller trigger is let up or the controller is not aimed towards the object, the animation deactivates. Once the animation path is complete, a sound clip is played to cue the end and another light in the scene turns on, illuminating a second object. This cascade is repeated: as each animated object reaches a destined position, a new object is illuminated and triggered with the viewer controller. When the final object in a given series completes its animation, a doorway appears, and the viewer moves into the passage using movement controls on the hand-held controller, triggering a scene transition to a 360-degree image space where a system clock is instantiated. A sequence of panoramic images cycles and displays around the viewer, and a voice-over narration is played. When the system clock reaches a given duration, a scene transition is triggered and the viewer returns to a game-space scene, confronted by a lit object to be animated. This

sequence of scene transitions repeats through three photographic and interactive scenes with an additional prelude and finale scene at the head and tail of the work.

Object interaction and viewer movement in the work are implemented by mapping programmed events to inputs on the controller device. The audience learns the controls with an instructional scene that opens the piece consisting of a voice-over description and demonstration of controller operations. Movement in the space is restricted to the horizonal axis (forward, back, right, and left) and is controlled by pressing various quadrants on the controller trackpad (Figure 41).

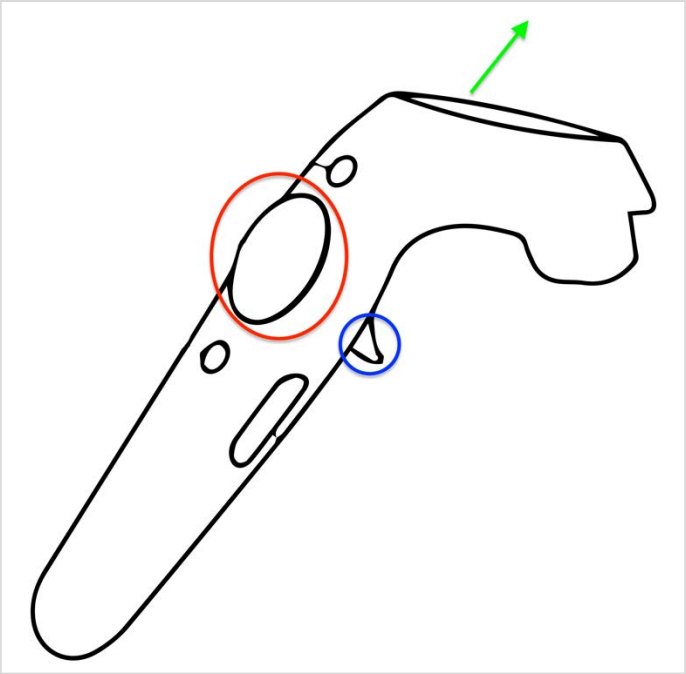

Figure 40: hand-held controller: trackpad (red), trigger (blue), ray-casting (green)

This horizontal motion is relative to the forward-facing direction of the viewer, so for instance, if the right-most area of the trackpad is pressed, the viewer will move in a direction ninety-degrees clockwise from where they are looking. This lateral movement

feature was added for more agility after testers had difficulty positioning themselves through doorways. The speed of the viewer's displacement is approximately 0.25 meters per second, significantly lower than the average walking speed. This reduced speed combined with forward orientation movement is designed to counter simulator sickness, an effect common in virtual reality experiences caused by a disconnect between the visual and vestibular system. Viewers can freely navigate and interact with the spaces at their chosen pace.

In addition to the trackpad movement control, the viewer can activate animations on select objects by pointing the controller and pressing the trigger button on the device (Figure 41). First, the objects to be triggered are chosen according to their proximity and visibility to the initial viewer position in the scene. For instance, in the third game-space, dozens of moving objects make up the set, but only nine are available to be triggered (Figure 42).

Figure 41: Game-space: interactable objects with collision detection component (green cubes) are chosen in relation to initial viewer position (orange square)

A communication network is established between the viewer controller and the animated object. Collision components (Figure 42) are attached to the objects and a virtual ray is cast forward from the top of the controller into the scene space. If contact is detected between the trajectory and the collider, the animation is activated, and the object moves (Figure 43).

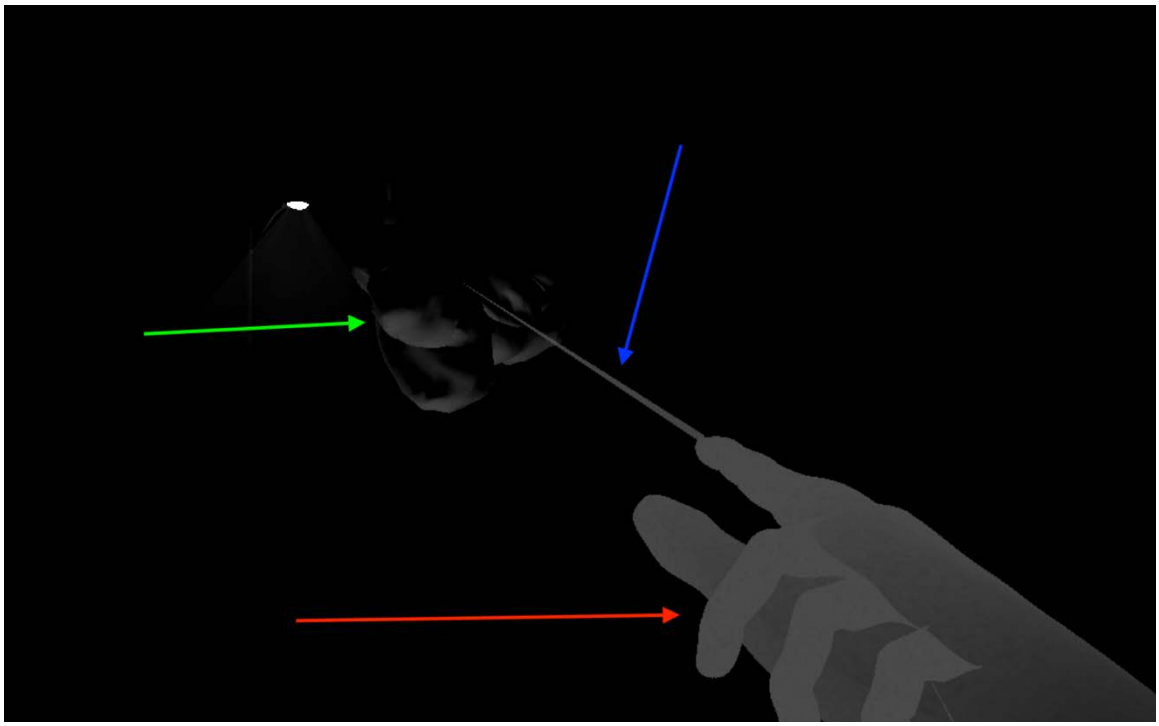

Figure 42: Screen captured shot with controller (red), interactable object(green), virtual ray (blue)

Since the audience task is to rebuild a three-dimensional structure with a fixed shape, the objects must be moved and positioned in a specific order. A global scene-wide condition system is implemented to ensure a linear sequence, so for instance, object 2 cannot be activated unless object 1 has completed its animation and so forth (Figure 44). In addition, the condition system dictates lighting cues that direct viewer attention toward the focal object at any given time.

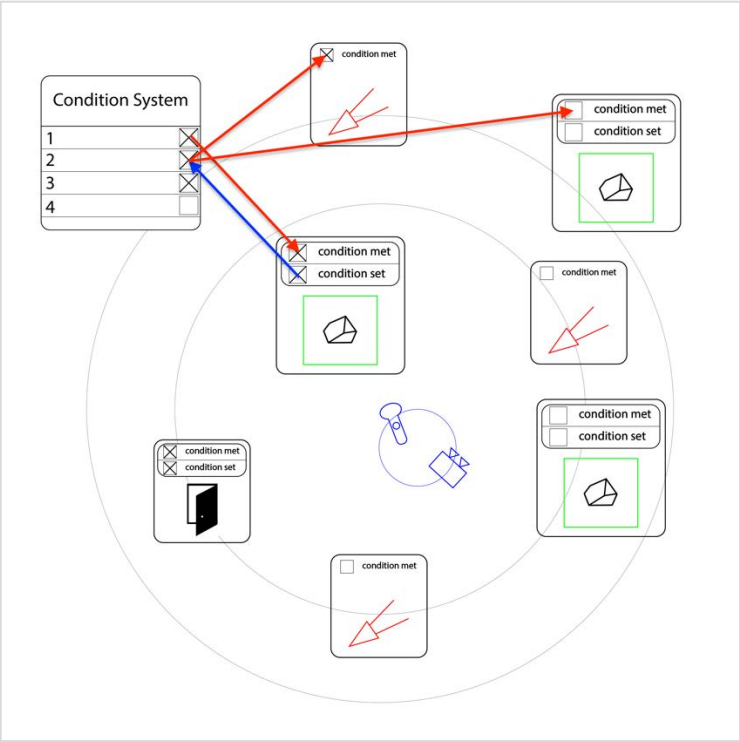

Figure 43: System Diagram: condition system object stores the status of objects to maintain sequential order.

Code segment example from above system:

```
void Update () {

    float  nTime = anim.GetCurrentAnimatorStateInfo(0).normalizedTime;

    if (conditionObject.GetComponent<AnimCondition>().animConditions[condMet])
    {
        GetComponent<Collider>().enabled = true;
    }

    if (nTime >= 1.0f)
    {
        if (!endSoundHasPlayed) {
            endSound.Play();

            endSoundHasPlayed = true;
        }
    }

    if (isPlaying && conditionObject.GetComponent<AnimCondition> ().animConditions[condMet]) {
        if (nTime < 1.0f) {

            anim.SetFloat ("speed", speed);

            if (!soundHasPlayed) {

                pieceSound.Play();
                soundHasPlayed = true;
            }
        } else {

            anim.SetFloat ("speed", 0);
            conditionObject.GetComponent<AnimCondition> ().animConditions [condSet] = true;
            soundHasPlayed = false;
            GetComponent<Collider>().enabled = false;
        }
```

If the condition number assigned to this object is met, continue

enable detection from the controller virtual ray

if the animation as finished (or reached normalized value of 1), play the ending sound cue

If the virtual ray is detected and condition is met, continue

if the animation clip is not finished, set its speed and play the object sound clip once

when the animation clip finishes, set the condition number assigned to this object and disable the virtual ray detection

As viewers progress through the scene animations, lights are activated with the same conditional system to illuminate the next task. Similarly, the ambient lighting for the background content increases with each scene so the visual depth of the spaces expands as audiences go deeper into the work. With the exception of these environmental lights, illumination originates from referent objects in the scene such as streetlights, work lights, flashlights, a lamp etc. in order to maintain a consistent narrative logic within the scenes (Figure 45).

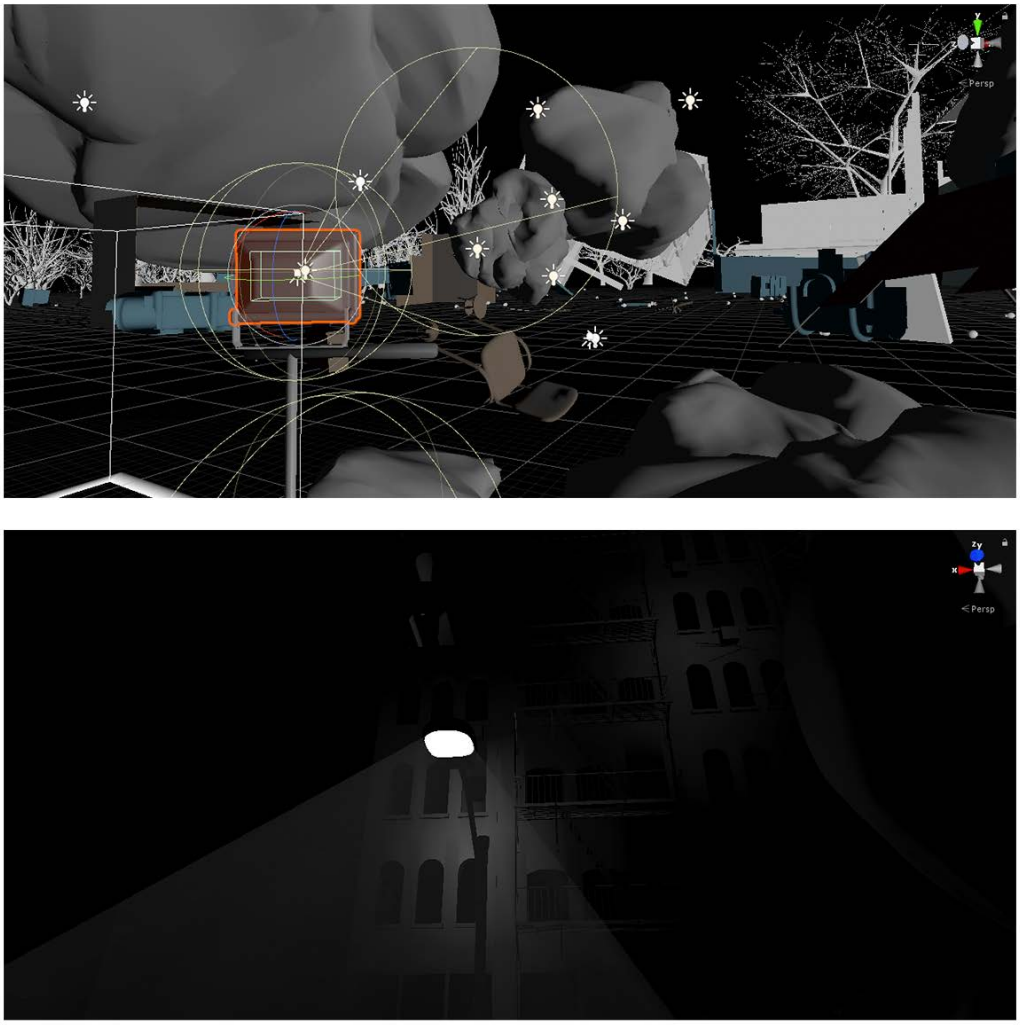

Figure 44: lighting is arranged to guide viewer attention (top) screenshot from piece (bottom)

Section 4: Conclusion / Preclusion

The Final Image is hosted online where the content can be accessed and installed without cost, and currently, it is being processed and formatted for additional distribution to broaden the exposure and ultimately engage more people. Similarly, the work is being entered into emerging media categories in several film festivals. The project is a living document, and in technical / industry terms, it is a beta software program that will be debugged and improved for the next year due to its complex nature. For instance, a couple of viewers have not been able to activate animations on objects when in close proximity to them. This is a design flaw overlooked in development and will need to be resolved and updated. Viewers have also strayed away from the active areas in scenes and become lost, disoriented, and unable to find their way back. A decision has to be made if this is acceptable, or whether visual cues should be added to help them establish bearings in the dark spaces.

In addition, *The Final Image* will exhibit in a physical art gallery in 2019 posing a new set of spatial considerations for the work. For one, tension arises when the individual nature of virtual reality display is inserted into public space. A quiet zone will be sectioned in the gallery where viewers can enter the experience with minimum interference. Further, the visual, sonic, and textual materials will be modular and dispersed in the space to be experienced as separate and relational components of a system. Prints of completed spherical image spaces and individual pictures will be framed and mounted, text content and maps of the 3D spaces will be displayed, and audio stations will provide extended unedited versions of the monologues for listening. The

compositional and institutional problems that arise from adapting the work for exhibition are part of the ongoing research and discovery process for the project and will inform decisions in future project work.

The Final image is a beginning, a foundation and springboard for future work. Similar to A Mountain in Two Eyes, or Memoir II, the project exists in a lineage and will spawn new ideas, experiments, and productions. This document proposes an immersion defined as spatial resolution fused with choice, a merging of environment and consciousness that envelops the internal and external world of a viewer. This idea is timeless and agnostic of medium, and the theory that supports it will evolve. *The Final Image* artwork is a manifestation of ideas, an instance, an experience implemented through sounds, images, and mechanics. This document and the materials and structures that compose the project are an archive, a snapshot of the present condition, and hopefully, will serve future research from its home in the library system.

Bibliography:

Arnheim, Rudolf. *Film as Art*. ACLS Humanities E-Book. Berkeley: University of California Press, 1957.

Bazin, André, Andrew, Dudley, Gray, Hugh, and Renoir, Jean. *What Is Cinema?. Volume 1*. Berkeley: University of California Press, 2005.

Bioy Casares, Adolfo., and Simms, Ruth L. C. *The Invention of Morel*. New York Review Books Classics. New York: New York Review Books, 2003.

Cortázar, Julio., and Blackburn, Paul. *Blow-up, and Other Stories*. 1st Pantheon Pbk. ed. Pantheon Modern Writers. New York: Pantheon Books, 1985.

Cremin, Colin/Ciara. *Exploring Videogames with Deleuze and Guattari: Towards an Affective Theory of Form*. Taylor and Francis, 2015.

Cremin, Colin/Ciara. *The Formal Qualities of the Video Game: An Exploration of Super Mario Galaxy with Gilles Deleuze. Games and Culture* 7, no. 1 (2012): 72-86.

Curry, Jessica. *The Music of Dear Esther: Creating Powerful Scores with Limited Resources.* Conference Talk. Game Developers Conference (GDC), San Francisco, CA, 2013.

Dargis, Manohla. "And Now for Something Completely Meditative." *The New York Times,*2007, Late (East Coast).

Deleuze, Gilles, and Guattari, Félix. *A Thousand Plateaus: Capitalism and Schizophrenia*. Minneapolis: University of Minnesota Press, 1987.

Eisenstein, Sergei. *Film Form: [and] the Film Sense; Two Complete and Unabridged Works*. Vol. MG 10; MG 10.;. New York: Meridian Books, 1957.

Fullerton, Tracy. *Breaking the Mold: Experiments in Evolving Game Narrative from the USC Game Innovation Lab.* Conference Talk. Game Developers Conference (GDC), San Francisco, CA, 2014.

Fullerton, Tracy. *Game Design Workshop: A Playcentric Approach to Creating Innovative Games, Fourth Edition*. 4th ed. Milton: Chapman and Hall/CRC, 2018.

Koerner, Joseph Leo. *Caspar David Friedrich and the Subject of Landscape*. 2nd ed. London: Reaktion, 2009.

Merleau-Ponty, Maurice, and Edie, James M. *The Primacy of Perception: And Other Essays on Phenomenological Psychology, the Philosophy of Art, History, and Politics*.

Northwestern University Studies in Phenomenology & Existential Philosophy. Evanston, Ill.]: Northwestern University Press, 1964.

Marks, Laura U. *The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses*. E-Duke Books Scholarly Collection. Durham: Duke University Press, 2000.

Marks, Laura U. *Immersed in the Single Channel: Experimental Media from Theater to Gallery*. *Millennium Film Journal* no. 55 (2012): 14-23.

Nabokov, Vladimir Vladimirovich. *Pale Fire*. First Vintage International ed. Vintage International (Series). New York: Vintage Books, a Division of Random House, 1989.

Quasha, George, Hill, Gary, Stein, Charles, and Cornwell, Regina. *Tall Ships*. Gary Hill's Projective Installations; No. 2. Barrytown, New York: Station Hill Arts, 1997.

Rutledge, Virginia. "Art at the End of the Optical Age." *Art in America* 86, no. 3 (1998): 70-77.

Tarkovsky Andrei. *Sculpting in Time: Reflections on the Cinema*. London: Bodley Head, 1986.

Wachtel, Edward. *The First Picture Show: Cinematic Aspects of Cave Art*. *Leonardo*26, no. 2 (1993): 135-40.

Wood, Michael. *The Magician's Doubts: Nabokov and the Risks of Fiction*. Princeton, N.J.: Princeton University Press, 1995.