Embodied Montage:
Reconsidering Immediacy in Virtual Reality

by

Halil Deniz Tortum

B.A. Film & Electronic Arts, Bard College, 2012

SUBMITTED TO THE
PROGRAM IN COMPARATIVE MEDIA STUDIES/WRITING
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN COMPARATIVE MEDIA STUDIES
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2016

Signature of Author:_____
Department of Comparative Media Studies/Writing
August 12, 2016

Certified by:_____
William Uricchio
Professor, Comparative Media Studies
Thesis Supervisor

Accepted by:_____
Heather Hendershot
Director of Graduate Studies, Comparative Media Studies

Embodied Montage:
Reconsidering Immediacy in Virtual Reality

by

Halil Deniz Tortum

Submitted to the Department of Comparative Media Studies/Writing
On August 12, 2016 in partial fulfillment of the requirements for the degree of
Master of Science in Comparative Media Studies.

Abstract

Virtual reality, as the name implies, is implicitly evaluated by its ability to reproduce or imitate aspects of reality, particularly in ways that convince users that they are physically present in the virtual space. This approach to the medium, a form of immediacy, effaces the material reality of the medium and can obscure the ways through which a medium can offer new forms of knowledge. Virtual reality's realism is a construct, and by acknowledging this, creators can push the medium to more experimental and novel ends. The framework of *embodied montage* proposed in this thesis seeks to provide an expressive vocabulary and techniques for creating virtual reality work that accounts for the material aspects of the medium.

Drawing on theoretical research from film and media theory, analyzing existing virtual reality work, and discussing the process of creating an original virtual reality work, I analyze the 3D capture and interaction design process. I offer the term *machine vision perspective* to describe the process of experiencing such images through virtual reality systems. This process provides a ground to experiment with embodied modes of thinking that are not possible with the human body situated in the real world. This makes *embodied montage*, a novel framework proposed by this thesis, possible. Embodied montage is the decoupling and recoupling of action and perception in virtual reality experiences in order to generate new meanings, similar to montage in film. Drawing on film theory, media history and cognitive science, this framework creates a territory for creative expression that challenges embodied cognitive structures and could effectively use the medium in ways that are distinguished from other media forms.

Thesis Supervisors
William Uricchio, Professor, Comparative Media Studies
D. Fox Harrell, Professor, Comparative Media Studies

Acknowledgements:

I'm deeply thankful to my advisers William Uricchio and D. Fox Harrell for their invaluable mentorship throughout this process.

Thanks to Beril Binoğul, Doğan Tekin, Cerrahpaşa Tıp Fakültesi, Council for the Arts at MIT for making *Hospital with one entrance and two exits* possible.

In the course of writing this thesis, I had the chance to meet influential practitioners of virtual reality. Thanks to Oscar Raby, Karim Ben Khelifa, Michael Naimark, Rus Gant, Rachel Rossin, Alexander Porter, Marcos Novak, Ersinhan Ersin for your inspiring work and generous insights.

MIT Open Documentary Lab has been a place for me to constantly rethink new media and non-fiction. Thanks to all the members for your curiosity and expertise and thanks to Sarah Wolozin for encouraging me to explore my interests.

Thanks to my fellow CMS classmates Beyza, Lilia, Andy, Lacey, Kyrie, Lily, Anika, Gordon, for making my last two years significant.

Thanks to Beyza Boyacıoğlu, Çağrı Hakan Zaman and Nil Tuzcu for your excitement and for sharing your tremendous knowledge.

The conversations and experiments at our art collective shaped my thinking and clarified my ideas. Thanks to Alan Kwan, Julia Litman-Cleper, Nate Saucier for dadasciencing with me.

Thanks to my parents for your love and support even though you don't necessarily understand what I'm doing. I promise to explain all this one day. But besides everything, thank you for your friendship.

Thanks to Ainsley Sutherland for your endless camaraderie, collaboration and affection, and many other things I'm inadequate to articulate.

Contents

List of Figures

Introduction

"When I look for the lightning, it never strikes. When I

look away, it does."

(Marlene Kos, *Lightning)*

In a 1976 video piece by Marlene and Paul Kos, lightning strikes only when Marlene Kos looks towards the camera. When she looks back, it doesn't strike. Kos proclaims what is happening even though she cannot see positive confirmation and can only see that when she does look, the lightning doesn't strike. Is she causing the pattern or only sensing it? The connection between Kos' statements, her gaze, and the lightning is notable because it is not edited. There is no cut from Kos to the lightning—the piece is one long shot. Using the tightly-knit system of a car, camera, observer, and environment, *Lightning* creates an uncanny proposition: we cause events with our perception and we can know this through the use of a camera. We see a connection between the subject and the environment that we know to be impossible. This thesis approaches virtual reality from a perspective suggested by *Lightning*: the technical apparatus of virtual reality creates a territory for creative expression that is rooted in ongoing computational interaction with perceptual mechanisms. In this thesis, I draw on film theory, media history, and cognitive science, to argue for a framework for the creation and analysis of virtual reality work called *embodied montage*.

Figure i.i: A screenshot from the video art piece *Lightning*, by Marlene Kos and Paul Kos

Virtual reality, as the name implies, is defined and implicitly evaluated by its ability to reproduce or imitate aspects of reality, particularly in ways that convince users of "presence," or a belief that they are physically present in the virtual space. Often, this means striving to use the user's perception and the points at which a user connects to the technical apparatus *seamlessly*, so that these points of contact are, ideally, unnoticeable. This thesis opens with an exploration of how this approach to the medium, namely works that aim for *immediacy*[1], can be limiting, both politically and artistically. Then, I discuss the capture and display technologies involved in virtual reality work. I analyze the 3D capture process, through which spatial aspects of the real world are captured for virtual reality and I offer the term *machine vision perspective* for experiencing such images through virtual reality systems. Then I propose the term *embodied montage* as an expressive

---

[1] *Immediacy* is a concept by media theorists Jay David Bolter and Richard Grusin. It defines a visual representation style "whose goal is to make the viewer forget the presence of the medium" (272). See Chapter 1 for a detailed discussion.

technique for creating meaning in virtual reality work. Virtual reality creates new relationships between action and perception. The theory of embodied montage allows us to break with realism and immediacy in analyzing how virtual reality work can modify these new relationships.

I argue that certain types of realism in virtual reality that emphasize forms of perception mimetic to real world experience, and are often considered desirable, are in fact similar to continuity editing in the classical cinema[2]. My thesis suggests that by breaking the illusion of naturalistic realism, particularly within the body, we are able to create new and additional meanings from adjacent sensations. For example, we expect certain outcomes from our body's movement and perception –like looking completely around to see what is behind us. That action isn't necessarily a given in a virtual reality environment and can be altered expressively. In this sense– that I am focused on the resonance between formal qualities of the medium and how they produce meaning– my approach is similar to that of the twentieth century modernists. Embodied montage originates in a study of film history, specifically the history of montage. I also hope this work guides the film audience in navigating the territory of virtual reality.

My methodology in this thesis combines critical analysis with critical making, situated between theory and practice. I draw on the theories of media scholars and analyze recent non-fiction work by virtual reality creators. Additionally, I reflect on the process of creating a virtual reality piece, *Hospital with one entrance and two exits.* In addition to being a stand-alone creative work, this project gave me applied experience both with the capture technology and the interaction design process through a game engine, giving me further insight on the mechanisms of technologies surrounding virtual reality. My experience with this approach evolved into the articulation and illustration of the affordances and characteristics of virtual reality images and of the *embodied montage* technique.

---

[2] See Chapter 3 for detailed discussion.

Chapter 1 discusses the concept of immediacy and the use of a medium to mimic existing human perception. Drawing on examples and theories from media and film history, I argue that immediacy in media obscures the expressive potential of a medium to expand human perception. I present examples from media history that argue that new media technologies make new forms of knowledge possible. The chapter ends by arguing that the realism in virtual reality is constructed and in order to understand the creative potential of the medium we need to analyze the process of this construction.

Chapter 2 introduces a real-time 3D virtual reality piece, *Hospital with one entrance and two exits*. I discuss the process and materials involved in this project in order to both introduce the reader to the affordances of the virtual reality image and procedures and technical constraints of a virtual reality project. I discuss two categorical aspects of the images that I derive from both the theoretical framework of the "softimage"[3], i.e., the image as programmable data proposed by media theorists Ingrid Hoelzl and Remi Marie, and from my process of working on *Hospital with one entrance and two exits*: the data set, and the rule set. Then I offer the term *machine vision perspective* for the process of experiencing such images through a virtual reality system.[4] Together, these two components suggest how the medium of virtual reality, and softimages, can shape our epistemological values.

In chapter 3, I focus on the experiential design of *Hospital with one entrance and two exits*, as well as a variety of non-fiction[5] virtual reality projects created by others chosen because of how they break away from naturalistic realism. I look at the affordances of these images when combined

---

[3] See Chapter 2 for detailed discussion.

[4] Currently, a virtual reality system generally includes a head-mounted display, peripherals (such as hand controllers) and motion trackers. See Chapter 2 for detailed discussion.

[5] With non-fiction, I refer to the works that are rooted in the real and try to expressively portray reality. I prefer the term non-fiction instead of documentary, because I am including not only documentary works but works that come from experimental media practice and contemporary art as well.

with embodied viewing technologies, such as the head mounted display. I explore how a reactive virtual space that builds new connections between action and perception not available in "reality" can be examined. In this chapter, I propose *embodied montage*: the use of a viewer's physical perceptual system and body in the creation of a montage. I situate this analysis against the film history and use of the montage technique.

Chapter 1:

Reconsidering Immediacy

"You'll feel like you're really there."

(Oculus.com, 2016)

2.1. The Problem with Immediacy

In an interview with *Wired* magazine, Palmer Luckey, the inventor of the Oculus Rift virtual reality headset and the founder of tech company Oculus VR, and John Carmack, the company's CTO, give some insights into how the business leaders of the industry envision the future of virtual reality.[6] Carmack states that virtual reality can improve the lives of many people, especially those living in poverty.[7] He underlines that Oculus and comparable technologies are relatively affordable, which for him means that "some fraction of the desirable experiences of the wealthy can be synthesized and replicated for a much broader range of people" (Au). The poor can live a life of opulence in a virtual world, without any incentive to fight against global inequalities. In such a scenario, the problems in the world become virtually non-existent. Luckey openly shares his ideas about how to build an ideal virtual future: "Once you've perfected virtual reality, you can imagine a world where you don't need to perfect anything else" (Au). What would this mean in today's world? Perhaps the problems around immigration, war, global income inequality would swiftly disappear. Governments, instead of attempting to integrate masses of refugees into society, could ship truckloads of virtual reality headsets to refugee camps. For Carmack at least: "If people are having a virtually happy life, they are having a happy life. Period" (Au). The social and political connotations of assumptions like Carmack's are grave, and reinforce notions of technological solutionism[8] that also constrain virtual reality work to a type of "naturalistic realism": the better the illusion, the better one can imagine one's real-world needs being met virtually, the more successful

---

[6] Virtual reality has had a variety of definitions since its inception as a term. In this thesis, I am using it to refer specifically to virtual reality work, which uses a head-mounted display and permits some interactivity with a virtual space. This is in contrast to computer based virtual worlds, such as the popular game Second Life, where there is no headset or peripheral technologies that take advantage of a user's perceptual system to allow her to navigate in the virtual space.

[7] This is an example of *technological determinism*, a theory which holds that technology drives social change: this is in contrast to theories of mutual reinforcement and shaping, where social concerns respond to and shape technology at the same time (Smith and Marx IX).

[8] Technological solutionism is the belief that technology drives solutions to problems, absent social, political, economic, and other forces.

the technology has been. This pattern of thinking is illustrated by film theorist Vivian Sobchack, in writing about the turn towards virtuality and an existence mediated by virtual images:

> Such an insubstantial electronic presence can ignore AIDS, homelessness, hunger, torture, the bloody consequences of war, and all the other ills the flesh is heir to outside the image and the datascape. It can ignore the lived body that not only once imagined its techno-logic but gave it substantial grounding, gravity, and value. It can ignore its own history. Indeed, devaluing the physically lived body and the concrete materiality of the world, the dominant cultural and techno-logic informing our contemporary electronic "presence" suggests that—if we do not take great care—we are all in danger of soon becoming merely ghosts in the machine. (Sobchack)

This passage, and the rhetoric surrounding the potential prosocial abilities of virtual reality described by Carmack and Luckey, point towards a complex relationship to "the real" embedded in virtual *reality*. The *realism* of virtual reality is founded on how convincing an illusion it provides: this is tied to decreased latency, refined graphics, and convincing movement modalities. In short, how well it deceives a user into believing that she is somewhere else. In this thesis, I will refer to this type of "real" as "realism" or "realistic." Aside from the techno-social dystopia implied by the prevailing rhetoric around virtual reality, as illustrated by Sobchack, this prioritization of realism at the expense of the real also greatly limits the scope of what is possible with virtual reality.

These assumptions are not only inherent in how the Oculus founder and CTO view the technology; they are also at the center of many discussions about virtual reality. While these assumptions might be attributed to commercial hyperbole, the hyperbole of the market itself actively shapes production, tools, topic and funding distribution and as such the rhetoric has a real impact on how the technology is used. Chris Milk, a public leader in the field of virtual reality, calls

the technology "an empathy machine." In his TED talk, he describes what virtual reality can do: "...

when you're sitting there in her [Sidra's] room, watching her, you're not watching it through a

television screen, you're not watching it through a window, you're sitting there with her" (Milk).

What is reflected in this sentence is a popular idea around virtual reality, namely "presence." Terms

like "presence" and "being there" are emphasized often in talks, writing and research around virtual

reality.

The idea of "presence"[9] in virtual reality is similar to what media theorists Jay David Bolter

and Richard Grusin call *immediacy*: "[a] style of visual representation whose goal is to make the

viewer forget the presence of the medium (canvas, photographic film, cinema, and so on) and

believe that he is *in the presence* [emphasis added] of the objects of representation" (272-273). The

medium disappears.[10] The distance between the original object of representation and the mediated

image is collapsed making the mediation through technology invisible. There is an ongoing desire

for immediacy in the history of media technologies from linear-perspective painting to

photographic art. This desire is "the desire to get beyond the medium to the objects of

representation themselves" (83). According to Bolter and Grusin, virtual reality gets closest: "in

virtual reality the viewer steps through Alberti's window and is placed among the objects of

representation" (83).

However, this *immediacy*, which renders the medium transparent, also constrains it. It limits

creative potential.[11] Limited in this fashion, the medium is also less able to potentially provide new

forms of knowledge by *shifting* perception or changing it. In the next section, I discuss how media

---

[9] The term "presence" has many different purposes and definitions. For one of the alternate and complicated views on "presence", see Hans Ulrich Gumbrecht's *Production of Presence.*

[10] Chris Milk, in an interview, says: "Ultimately, what we're talking about is a medium that disappears, because there is no rectangle on the wall, and there is no page you're holding in your hand. It feels like real life" (Johnson).

[11] Many virtual reality works recreate the human perspective and real world physics. For example, it is rare to see a world upside-down in a virtual reality piece. This somewhat parallels early cinema works which recreate the experience of watching a theater play. In both instances, there is the idea of a viewing subject which does not necessarily fit the new expectations of the medium.

technologies have a reciprocal relationship to the production of knowledge, through the ways they break away from realism, and argue the importance of dissociating virtual reality technology from the desire for immediacy.

1.2 Media Technologies and Production of Knowledge

Walter Benjamin coined the term 'optical unconscious' to address the potential of the photographic medium (37). For Benjamin, a new means of perception comes into existence with the apparatus itself. The photographic medium can offer insights into layers of the perceptible world that we were not aware of before. For example, a development such as slow motion "discloses quite unknown aspects" of movement (37). Eadweard Muybridge, in his principal experiments with the moving image, photographed a galloping horse in *Sallie Gardner at a Gallop* to understand how photography can change the range of perception. In this experiment, the formerly imperceptible motion of the horse was revealed: all the four legs leave the ground at one point during a gallop. The microtime and the non-perceptual temporality made possible by photography was a scientific trope in the 19th century[12]; "photography was superior to direct observation" because it captured events taking place at intervals shorter than a human eye can perceive (Canales 118). As discussed in great detail in Lorraine Daston and Peter Galison's *Objectivity*, photography promised a view of the world "uncontaminated by interpretation" and created a new epistemological claim, namely mechanical objectivity, that caused a fierce debate in scientific fields (139). Was it the human eye or the machine that saw the world in its true form? Both had limitations but it was obvious that the camera offered a new form of perception. Another instance of the apparatus co-creating with the human observer new forms of perception is the close-up. A close-up shot expands space and "brings to light entirely new structures of matter" (Benjamin 37). An expression of a face or a detail

---

[12] Jimena Canales presents the invention of microtime in her book *A Tenth of a Second.*

of an object is suddenly revealed. With photographic media, the familiar becomes strange and formerly imperceptible details of objects, actions, expressions are made clear on a large projection. The camera is perceived as the ideal observer, "patient, indefatigable, ever alert, probing beyond the limits of the human senses," producing "images that depicted the physical world in its full-blown complexity, its asymmetrical individuality…" (Daston and Galison 16, 139). Photographs quite literally allowed us to see the world differently.

Figure 1.1: All the four legs of the horse leave the ground in *Sallie Gardner at a Gallop*

Jonathan Crary, in his book *Techniques of the Observer* takes this idea further and implies that the invention of the camera obscura might have led to the modern subject-object relationship. With its clear spatial separation between the viewer and the object, the camera obscura creates an isolated, enclosed observer cut off from the exterior world; a relation which reflects a worldview based in body-mind dualism. The camera obscura helps shape modern Western thought and contributes to a shift away from the pre-modern organization of knowledge, creating "a free

sovereign individual and a privatized subject confined in a quasi-domestic space, cut off from a public exterior world" (Crary 39). Vision is dislocated from the body and the act of seeing is dissociated from the human consciousness. Instead of working in tandem with other senses of touch or smell, vision becomes its own separate entity. "A space informed by human consciousness" is replaced with a space that is informed by the apparatus, by the technology (Benjamin 37).
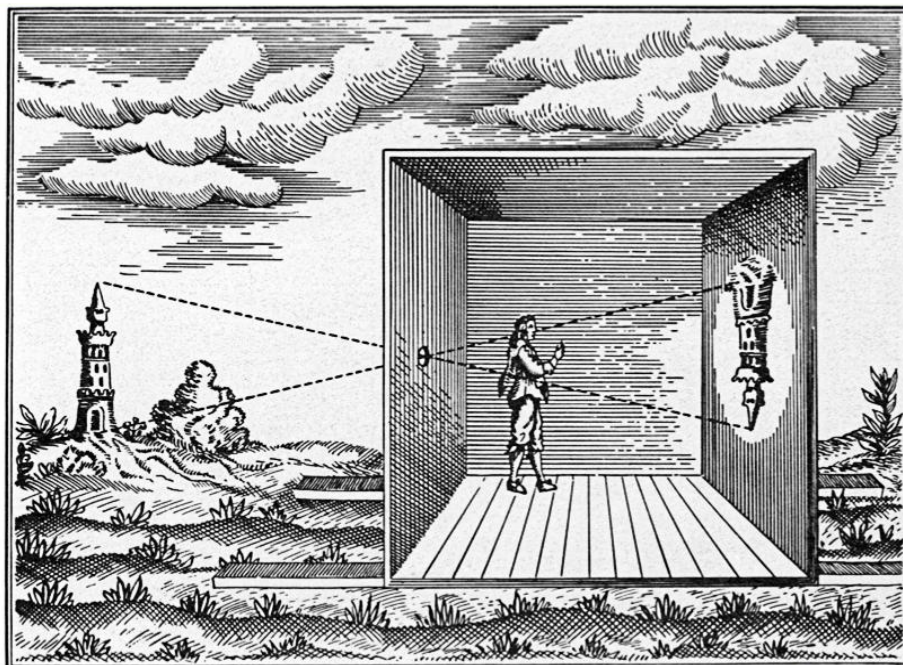
Figure 1.2: The subject of camera obscura

What Benjamin's and Crary's writings and the scientific debates of 19th century add to the discourse around media is a reverse logic. New technologies and media innovation are driven by emergence of novel knowledge through different factors, such as science and engineering research. However, new media has the generative potential to produce unrealized intellectual and perceptual realities. Though a plethora of technical research results and a large body of knowledge led to the invention of the cinematograph, cinematography itself then prompted new scientific questions regarding perception (Schroeter 28). New forms of media generate new forms of understanding. As

Benjamin points to with the concept of "optical unconsciousness" and Crary demonstrates with his study of the camera obscura, media technologies can change what is considered "knowledge," with unpredictable effects that go beyond scientific regimes (for Crary, the re-definition of the observing subject).

To recap, for both Benjamin and Crary: ***New media technologies help us to create new forms of knowledge,*** through the ways in which they change how we are able to perceive the world.

This lens has not yet been satisfactorily applied to virtual reality. The main discourse is that virtual reality is a technology to create an alternate reality, which is virtually indistinguishable from reality itself, as manifested in Cormack's statement about replicating the "desirable experiences of the wealthy" for other people. In this case, the technology is thought to provide an exact replica of the reality, offering new sights. However, it is unable to offer new ways of seeing and new forms of knowledge. And that is partly because of how we refer to the technology itself:

virtual *reality.*

This is the first medium that has a referent to reality in how it is described.

Cinema means "movement writing."

Photography means "light writing."

Phonography, "sound writing."

These familiar technologies are framed through a practice—namely, writing—inscription via the medium. Virtual reality, in contrast, has an epistemological claim. In the case of virtual

reality, the technology claims to provide access to reality by way of seamless realism. Virtual reality is hailed as a *medium*, yet unlike other media forms mentioned above, there is no reference to construction or practice (the "graphy," or implicit inclusion of kind of translation: the medium of transmission). This rhetorical positioning suppresses the analysis and creation of virtual reality work in several ways. First, it conceals the material practices involved in constructing virtual reality work. Secondly, by concealing the material practices involved in construction, the expressive potential of virtual reality is limited to work which appears realistic. In these early days of virtual reality as a publicly available expressive medium, I examine the positioning of virtual reality and realism from the perspective of Rudolf Arnheim's discussion of cinema and realism. Drawing on the notion that new media creates new forms of knowledge, my methodology is inspired by art theorist Rudolf Arnheim's approach of analyzing cinema. In his seminal work *Film as Art*, Arnheim argues against the period's dominant concern with realism, and claims that the ways cinema breaks away from reality are instrumental in its art. He analyzes techniques of cinema and points to aspects such as planar projection, which causes a reduction of depth, and nonlinear film editing, which breaks the chronological connection between shots and creates an absence of space-time continuum (Arnheim 8, 9, 20). In these ways, the medium moves away from realism and creates its own language. Arnheim argues that film's so-called "realism" is in fact a construction; and as long as one is going to construct, one should embrace the idea and construct expressively, rather than effacing the act of construction. Arnheim was writing at the early days of film theory, and his approach provides a model for examining a young technology and locating it as an art form. Virtual reality can be subjected to a similar examination, both as an artistic and intellectual project. Dissociating the technology from the metaphor of "reality" diversifies how we think about the medium and judge projects made within it.

Chapter 2:

Machine Vision Perspective

"Of course, it is not like an ordinary photograph; this is my latest invention. We shall live in this photograph forever. Imagine a stage on which our life during these seven days is acted out, complete in every detail. We are the actors. All our actions have been recorded."

(from *The Invention of Morel)*

$$d = \frac{c\phi}{4\pi f_m}.$$

Formula to calculate distance for *time of flight cameras*.

2.1. Hospital with one entrance and two exits

Figure 2.1: Image from *Hospital with one entrance and two exits*

In parallel to my research, I have created *Hospital with one entrance and two exits*,[13] an experimental non-fiction real-time 3D project viewed in a virtual reality system.[14] The piece is installed with a head-mounted display that has positional tracking capabilities, Google's Project Tango device in this case.[15] The viewer finds herself in a virtual hospital environment and is free to explore by walking around in the environment. The virtual environment is created from point-cloud data generated by a laser scan acquired at the Cerrahpaşa Hospital in Istanbul.

---

[13] I will refer to this project as *Hospital with one entrance* for brevity.

[14] "Real-time 3D" is intended to distinguish this type of work from 360-degree video, as well as refer to how images are constructed in real-time to accommodate the user's movement, in contrast to pre-rendered, static models.

[15] Positional tracking means that a user can walk or move around a space, and that movement is detected either internally by the display device or by cameras mounted in the space (as is the case with the HTC Vive, another contemporary head-mounted display).

Inside the headset these points appear to be suspended in a volumetric space and resolve, at a distance, into recognizable shapes: an operation room. The viewer is surrounded by a long corridor, stretcher and surgical equipment composed by these points. The buzzing sound of fluorescent lights fills the environment. The viewer begins in the corridor and is free to explore; she can walk into the operation room or walk through the corridor. The point cloud reacts to the viewer, both to her gaze and to her head and body movement. This means that as the viewer moves, the points constituting the hospital corridor slowly appear to open up, with more and more empty space between them, and eventually dissolve. The viewer seems to have a *view from nowhere*, because she is able to move out of the perspective from which the laser scanner was located when it captured the data, and thus skew the perspective. Yet at the same time she finds her body used as an instrument in the construction of the virtual image. By this I mean that the viewer's movement modifies the image, which appears on the headset screen: from a small motion of her head to the larger motion of her steps. As she walks, a black, empty hole is left in the initial position of the viewer's body.

Figure 2.2: Image from *Hospital with one entrance*

The viewer walks into the operation room. The fan begins working loudly. When she enters the room, her field of vision expands, making the room larger, an effect which can be slightly nauseating. The viewer can explore the room and look at each object (operating table, medicine cabinet, nebulizer, ventilator, surgery lights etc.) yet each time she tries to get closer to the objects, the points separate and objects fade away. Clear observation is only possible from a distance. After a short time spent in the operation room, the viewer starts to feel a tapping on her left eye, a disconcerting feeling. The tapping feels physical, even though she knows this is impossible. The tapping continues until she leaves the room.
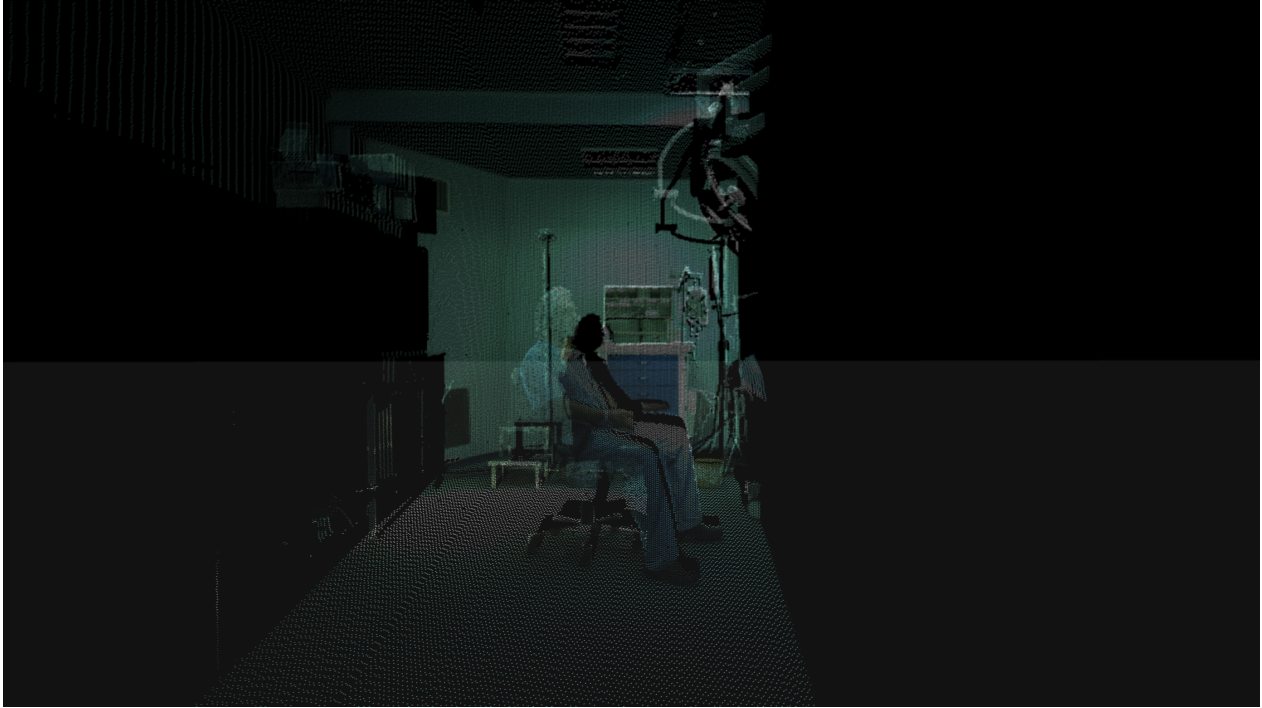
Figure 2.3: Image from *Hospital with one entrance*

At the other end of the corridor, there is another room. When the viewer enters this room, a new sound recording begins to play. There seems to be a doctor sitting in the room, though his form is, like the building itself, constructed from points that dematerialize as the viewer approaches. The sound recording is of the doctor recounting his experience of waking up from a surgery. This is the first time the viewer sees another human body visualized in a point cloud. The shadow of the doctor's already transparent body drops on the wall. The visibility of the doctor is determined by the viewer's movement, fading in and out of clarity depending on the viewer's position.

The piece ends in a new corridor full of stretchers. At this point, the constant sound of fluorescent lights humming is replaced by the blinking sound of broken fluorescent lights flickering. The rest of the building can be observed from this corridor, as everything is semi-transparent in the environment and the viewer can see through walls and objects. The viewer starts

walking towards the end of the corridor and points get sparse. The hospital turns into floating points in space.

Figures 2.4 & 2.5: Images from *Hospital with one entrance*

Hospital with one entrance is a project through which I will delve deeper into the material aspects of the production process of real-time 3D experiences and the images of virtual reality. In the rest of the chapter, I will describe the production of this piece and analyze the affordances of the images.

2.2. Softimage

In analyzing the capture and construction of these images, I use Ingrid Hoelzl and Remi Marie's theoretical framework of the "softimage." A softimage is "a programmable view of a database that is updated in real-time" (4). It is different from earlier types of images, namely a "hard image," as it is "ubiquitous, infinitely adaptable and adaptive and intrinsically merged with software" (7). Furthermore, "the image is not part of a programme" but rather it is a "programme in itself" (4).[16] The framework of the softimage allows one to look at images as software themselves.[17] It is possible to interact with the software through manipulating the images, and the images

---

[16] This is similar to Lev Manovich's principles of new media: numerical representation and modularity (27, 30).

[17] The softimage is more concerned with the changing photographic paradigm and representation of the world through the photographic perspective. Because of that it implicitly excludes computer generated images. Similarly, I will exclude CGI and only focus on 3D capture.

themselves should reveal characteristics of the contingent software. It is important to understand

that the images has properties, we may not perceive them all, but we need to start thinking of

images as a kind of responsive and contingent entity, capable of acting in complex ways.

Approaching the images of *Hospital with one entrance* through the softimage framework

allows for a constructive analysis. The two domains I will look at are "images as data set" and

"images as rule set." The data set refers to "data" that are obtained through 3D capture, measuring

spatial coordinates and color values of the points in an environment. The rule set refers to how code

is organized, then layered onto the virtual objects and spaces, and instructed to respond. After

introducing the technology of the laser scanner and the definition of "data set," I will explore four

aspects of the 3D images that emerge in relation to the data set: they are transplanar (between

planar and spatial), referential (have a real-world, indexical referent), virtual (non-optical) and

contingent (perspective, transparency and spatial relations of are subject to change). And, I will

discuss three aspects of the rule sets that govern the data: processual (the environment is created

continuously with the user's input), operative (the user can act on the environment) and embodied

(the user acts with her body with the use of virtual reality hardware). This categorization (figure

2.6) draws concepts from several media scholars –Hoelz & Marie, Uricchio, Sutherland, Schroeter,

Grau, Elsaesser, Manovich and Virilio– and combines them in an original framework. I will then

propose the term *machine vision perspective* to describe the generative interaction between

softimage and virtual reality system.

Image as Data Set	Transplanar (between planar and spatial)	Referential (have a real-world, indexical referent)	Virtual (non-optical)	Contingent (perspective, transparency, spatial relations are subject to change)
Image as Rule Set	Processual (the environment is created continuously with the user's input)	Operative (the user can act on the environment)	Embodied (the user acts with her body with the use of virtual reality hardware)	

Figure 2.6: Table detailing the affordances of the images

2.3. Image as Data Set

The hospital in *Hospital with one entrance* is a visualization of point-cloud data captured through a 3D laser scanner. For the laser scanning process of *Hospital with one entrance*, I worked with an architecture firm. Over the course of four hours, we created twelve different scans. The process started with placing the laser scanner at a location which will be the point of capture. The scan took approximately fourteen minutes: the instrument took seven minutes to collect coordinate data by emitting lasers, and seven minutes to take the photos of the place to gather color information and map RGB data to each individual point. Laser scanning (or 3D capture in general) is a unique form, distinct from previous forms of computerized media, such as digital photography. It records physical aspects of the place itself: spatial depth and proximities.

Many fields including architecture and engineering use laser scanning because of its ability to take precise measurements. The laser scanner uses a technique called range imaging, or more specifically time of flight technology, which "is based on measuring the time that light emitted by an illumination unit requires to travel to an object and back to the sensory array" (Kolb and Pece

51). A 3D laser scanner emits lasers in all directions and collects the coordinates of each point the

laser hits and returns. Thus it results in a spatialized point construction of a real place. A popular

manufacturer of laser scanners, a company called FARO, describes the capabilities of the device as

"[enabling] fast, straightforward and accurate measurements of...complex structures, crime

scenes...and production and supply facilities" with ±2mm accuracy ("FARO"). The device makes

the 3D capture of the physical world possible. Furthermore, the company argues that "high-

precision, point-by-point distance measurements" make the captured data highly reliable

("FARO").

Figure 2.7: 3D laser scanner capturing the OR corridor

Laser scanning is one form of 3D capture. In recent years, various techniques and

technologies that can accurately capture objects and scenes into 3D models have been developed.

These 3D capture methods produce images with "a high degree of geometric and photometric

fidelity." There is a wide range of such approaches. Some systems include "active range scanning"

such as laser-scanners or real-time depth cameras, such as Microsoft's Kinect. These instruments directly capture data from the environment. Another method of 3D capture is through "passive image-based modeling algorithms"; these algorithms take multiple RGB images and videos and algorithmically process these images to create 3D objects and environments. Examples for such systems include Autodesk's 123D Catch, Agisoft's Photoscan, or Microsoft's Photosynth web service (Szeliski IX). Each of these programs facilitates the modeling of real locations or objects in 3D space with indexical qualities, rather than purely rendered with no real-world data capture in 3D modeling programs.

Most of the graphics software that renders these images requires a solid surface representation in the process of creating a 3D environment or object. The point-cloud data is generally algorithmically processed to create a mesh reconstruction "'filling the holes' between the measured point samples" (Boubekeur 151). This process creates an approximation of the captured data through the algorithm and that approximation, in turn, is dependent on the code used in the reconstruction.

However, *Hospital with one entrance* arrests this process, and visualizes the captured point-cloud data without the mesh reconstruction. For the project, the point-cloud data was registered, combined and sampled using MeshLab and then imported into the Unity game engine to enable interactivity. This presentation of unmeshed point cloud data makes it possible to interact with a less-processed form of captured indexical data, which is a visual representation of a simple text file compiled of the coordinates of each point measured.[18]

The image comes as a text file: as a data set of coordinates. The file includes millions of coordinates with RGB color data (in the form of x,y,z and r,g,b). It contains depth and spatial

---

[18] There are other works that also do this, such as *In the Eyes of the Animal* and *Phantom*. See Chapter 3 for discussion of these works.

relationships. Thus it is a data set for potentially replicating the surface contours and spatial positioning information of objects and places that the device has captured. The image, in its essence, is not visual; the laser scanner captures machine-ready data of the world that can be visualized through computer software into what a human eye would perceive. This spatial data is presented as a 3D point cloud environment, arrested halfway between the text file readable by machines and a mesh which reproduces a more stable environment. I will analyze the image as data set as *transplanar, referential, virtual,* and *contingent.*

2.3.1. Transplanar

There are broad disciplinary differences in the definition of "3D images." A 3D image for a film scholar would likely mean stereoscopic photography, whereas for an architect it would mean computer generated 3D graphics. Media theorist Jens Schroeter underlines the scholarly reluctance to theorize about 3D images and states that they "represent a disturbing phenomenon" (37-38). This disturbance lies in the fact that 3D images blur "the line between image and space" and it is not clear whether they should be treated as images or "a visual phenomenon that is sui generis" (37). Most of the discussion around images is of planar images constituted through "projection of a three-dimensional object or several objects within space onto a two-dimensional plane" (33). Planar images depend on geometrical optics, on the assumption that the nature of light is linear rays. Geometrical optics make perspective possible through mathematical formulation. Through the introduction of perspective, images become *spatially* comprehensible. Planar images privilege a "planocentrism" through which images should be "readily comprehensible 'at a single glance'" (35).

Yet not all images are planar. What should we do with non-planar images such as "stereoscopy, holography, and certain forms of computer graphics" (37)? Schroeter underlines that

they are always categorized vaguely under the term "3D." This term does not explain the nature of such images, since because they do not occupy space in the physical world, they are not in fact "three-dimensional," they only appear to be so under certain conditions. Furthermore, all these images are based in different optical systems; stereoscopy in physiological optics, optical technologies depending on human eye and human cognition; holography in wave optics, an optical system based on the wave properties of light such as interference, diffraction, polarization; computer graphics in virtual optics, computational optical system able to recreate all the above-mentioned optical systems (27). Instead of the term 3D, Schroeter offers the term "transplane image" as 3D images are not planar but also not spatial (38). This nomenclature gives a better definition of the phenomenon as it emphasizes the inherent tension in 3D images between plane and space.

The images in *Hospital with one entrance* are transplanar and they employ physiological, wave and virtual optic systems. In terms of physiology, they are stereoscopic, each eye gets a slightly different image which creates the illusion of depth for the viewer. In this sense, they are presented on a plane, which is the screen, but they create an illusion of space through physiognomic optics. They are captured through a laser system, which is based on wave optics that can emit light coherently. They create a virtual space through virtual optics; perspective is algorithmic, and the viewer can walk in a real space and the images presented through the head-mounted display processually change according to the position and head movement of the viewer.[19]

2.3.2. Referential

Media historian Oliver Grau tries to pinpoint the properties of such images, which he calls "virtual reality's dynamic images" (248). Attempting to distinguish the ontology of these images in

---

[19] See 2.4.1. Processual for further discussion.

comparison to what used to be called "a picture," he argues: "unlike a photograph, for example, a computer image—an implosion of the image, the real, and the imaginary—is non-referential" (248). The computer image is non-referential because it does not refer to an actual, physical object. Grau says the lack of actual object is compensated for in how the computer image is structured recursively with appearance and code: "On the one hand, a property of the image is its visible appearance, which is concrete; on the other, its basis in numbers, or code, is an abstraction" (249). Grau's definition, however, doesn't account for all types of computer images. A computer image is both appearance and code at the same time; yet this distinction is true for anything digital, including digital photographs. Thus, Grau's definition clouds the different properties that exist between different type of digital images, such as a 3D scan, a parametric design object, and a digital photograph. Grau's theoretical approach to "virtual reality's dynamic images" is true in fact of all digital images, and hence does not provide us with a very productive structure to make sense of the interactive transplane image. In comparison, Schroeter's categorization based on spatiality and navigability seems much more fruitful. Furthermore, Grau's claim that a computer image is non-referential does not hold up to be true in all instances, such as in the case of *Hospital with one entrance*, or in the case of other images taken by the laser scanner, which as an assemblage of measurements are completely referential. These projects blend the distinction between indexicality and computer generated images.

Unlike virtually modeled objects or other type of computer generated imagery, these images have an indexical relationship to a place.[20] They are based on their registration of data from a physical space. Discursively, this new image is a new representation of reality. Companies working in this space make this claim regularly; software company Autodesk, calls the technology "reality

---

[20] Indexicality is a sign modality proposed by Charles Sanders Peirce. An index is a sign "whose relation to their objects consists in a correspondence in fact". An index has a causal relationship to the object it represents, thus signifying its existence. In this context, I am using the term in relation to film and photography theory (Bazin).

capture." Furthermore, the technology is being deployed to capture crime scenes and present these as evidence in court. The recent report on laser scanners put out by the National Institute of Justice's Forensic Technology Center of Excellence reads: "The data provide a completely objective analysis and highly credible evidence in a court of law ... Data obtained from scans document the entire scene and may provide spatial evidence first missed as relevant patterns or evidence not obviously visible" (Ropero-Miller 44). These use-cases for 3D captured images suggest exceptions to Grau's claim that the computer images are non-referential. Not only does a 3D laser scanner need to be physically in the place of capture in order to generate the data but it also produces a representation that is increasingly regarded as more accurate and reliable than photographic images. In the case of laser scanners, communities like Autodesk and legal courts serve as what Stanley Fish call "interpretive communities." They constitute the "properties" of texts and assign "their intentions" (483). They attach a meaning of *reality* to the laser scans. These images constitute a new representation of reality, with its own artifice.[21]

2.3.3. Virtual

However, it is important to note that such images are still virtual. Schroeter defines the "virtual image" as created through "mathematical formalization" and "algorithmic processing" in computers (377). The point-cloud hospital in *Hospital with one entrance* resembles a hospital, however, it is clear that this representation does not resemble the hospital the naked human eye would see it. The images in *Hospital with one entrance* offer a referential virtuality through machine vision. This is because the capture device is not optical in the sense of a camera; the lens is

---

[21] Regarding the "new representation of reality", media theorist Alexander Galloway's comment on data visualization resonates deeply: "It simply means that any visualization of data must invent an artificial set of translation rules that convert abstract number to semiotic sign. Hence it is not too juvenile to point out that any data visualization is first and foremost a visualization of the conversion rules themselves, and only secondarily a visualization of the raw data. Visualization wears its own artifice on its sleeve. And because of this, any data visualization will be first and foremost a theater for the logic of necessity that has been superimposed on the vast sea of contingent relations" (83).

not a surrogate for the eye. These images "cannot be captured on a screen (e.g. the mirror image)" (377); thus they are different than the 'real' images "defined by optics" (377). Unlike a photograph, which captures the light rays going through the lens and hitting a plane, a sensor or a negative, the 3D capture instrument captures machine-first data: the spatial coordinates of a place. Then, these coordinates are visualized through virtual optics for the human eye. Film historian Thomas Elsaesser separates such images from the "ocular-perceptual" ontology:

> the case of (photographic, analogue) cinema, we are operating within an ocular–perceptual frame of reference... Within this field, we are bound to see the virtual (as in VR) as also part of the ocular–perceptual register...but is itself not generated by or determined through the optical. Thus, in the digital, we do not have an ocular–perceptual ontology; instead we have a digital ontology. (300)

Laser scanning is machine vision visualization because it is machine-first data visualized. It is the conversion of a physical environment into a dataset instead of an optical image. It is a depth data collection translated to vision. It is not an ocular-perceptual ontology but a digital ontology translated into an ocular-perceptual epistemology. Thus it is different from digital photographs, which are based on an ocular-perceptual logic.

2.3.4. Contingent

The machinic image differs from optical images because it is virtual. Recall Schroeter's definition of virtual images as mathematically formalized and algorithmically processed. There is a "separation of structure/form and material substrate" at the core of these images; the form of the object is detached from the materiality of the objects (378). This quality makes these images infinitely malleable and contingent.

The image below looks like a photograph. However, the viewer can move away from the point of capture and skew the perspective. She can see outside of the apparatus' point of view. A phenomenon unique to this form of capture is the ability for the viewer to move away from the "point of capture," or the location of the laser scanner device. This is unlike photography or film, in that the perspective can be different from the point of capture (where a camera is located).
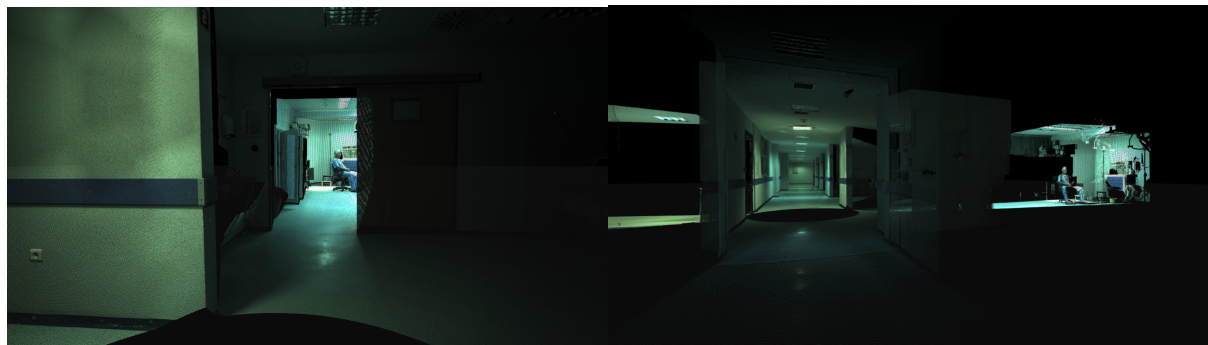
Figures 2.8 & 2.9: Moving out of the point of capture in *Hospital with one entrance*

The images are not non-perspectival, however the perspective is contingent; it can be skewed and the viewer can depart from the point where the original perspective is created. Prior traditions of images keep the body fixed: "from Renaissance monocular perspective to modern cinema, from Kepler's camera obscura to nineteenth century camera lucida, the body had to remain still" (Manovich 104).[22] However, in this case, there are different points of capture combined by the computer. In figure 2.9, the two black holes correspond to two different points of capture. The image is synthetic, point of view is not tied to point of capture. It is created through combining multiple captures into a single point cloud. It is a sum of multiple perspectives, all of which can be

---

[22] Manovich forgets the panorama, in which the body is not fixed. Although the panorama does not respond to viewer's movement similarly to a virtual reality experience, some aspects did not fit with Manovich's model: for example, a viewer can move about, choosing which part of the image to focus on, and cannot comprehend the entire piece from one position.

skewed by user movement. The points of view that exist in between point of captures are

synthesized by the computer at every refresh.
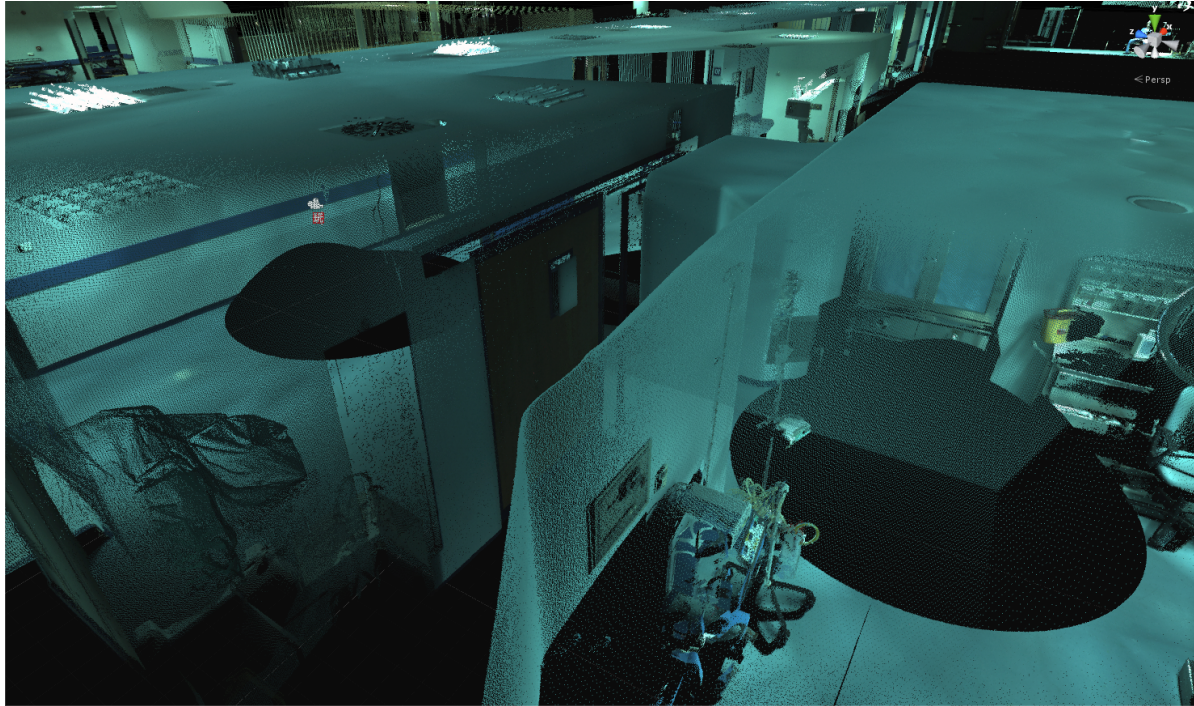
Figure 2.10: Image of the combined laser scans
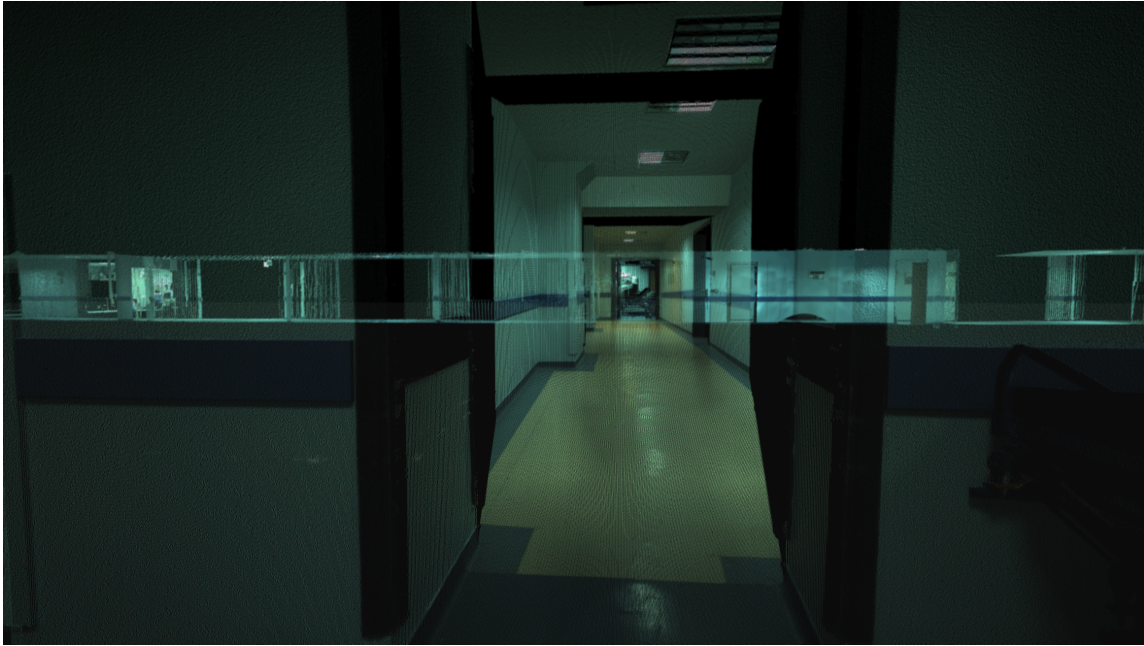
Figure 2.11: Image showing transparency from *Hospital with one entrance*

These images also have a transparent quality. Since the image is not optical– not created originally from a projection, not a set recording of light rays– but rather a data collection of the spatial coordinates of the place, the properties of the object, such as opacity, can be altered (the rule set). In an image, the opacity of the image itself could be altered, but not the objects and masses depicted within it. Previous image technologies have had similar impact, such as the x-ray. Architectural historian Beatriz Colomina describes another instance of a "transparent" image in her essay on the x-ray, and its effect on human perception. With x-ray photography the objects are penetrated and "the mysteries of the interior are brought to the surface by a screen." With the x-ray, "the whole world had to be seen again, or more precisely, as if it were a whole new world."

3D capture has similar potential. With the proliferation of Google Tango and similar devices (devices that uses real-time sensors and cameras to build meshes of a real environment, allowing apps to incorporate the real world into virtual reality and augmented reality experience), 3D images

drawn from our real environments become increasingly common. 3D capture turns the world into a kind of x-ray: once an object is scanned, it can be rearranged, moved, imbued with new properties. The captured place can be architecturally reconstructed. Any corridor can be connected to any other. Objects can be rearranged over and over. The image and the space are not fixed.
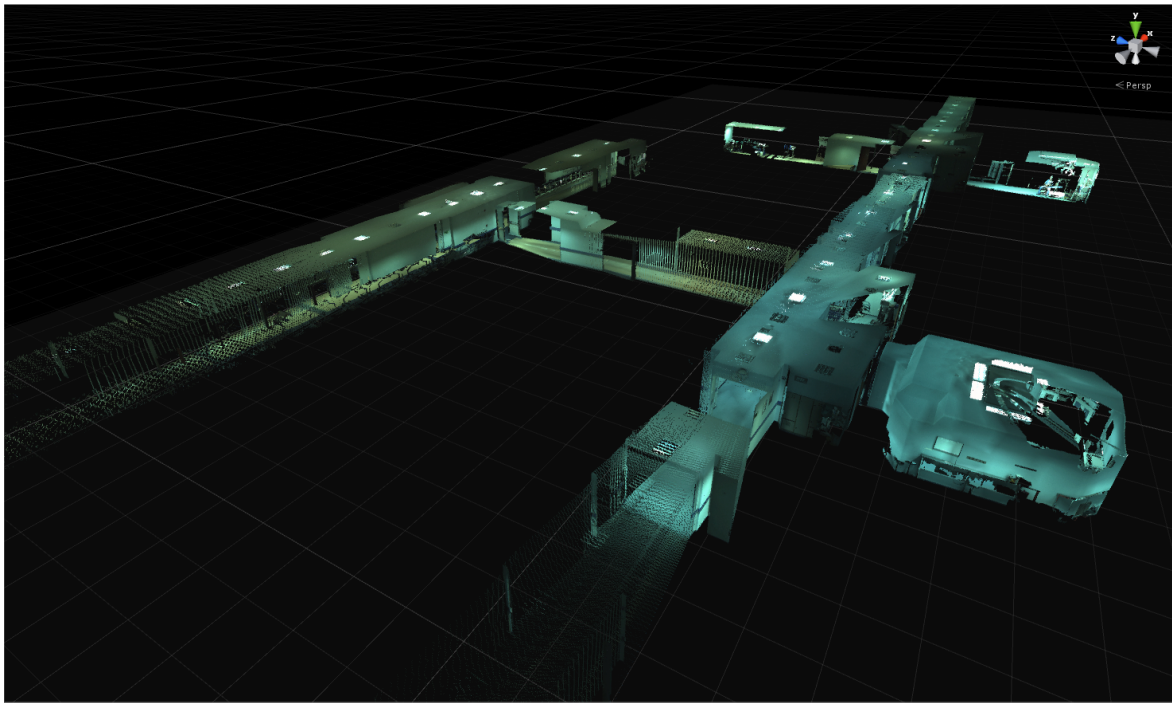
Figure 2.12: Image of the combined laser scans

2.4 Image as Rule Set

Hospital with one entrance is created using a Unity game engine. Most real-time 3D experiences are designed in game engines, tools originally intended to create computer games. In this environment, code can be attached to particular objects, imbuing them with additional characteristics. In Figure 2.13, we see the Unity game engine interface. In Unity, code can be embedded to objects, environments (the physics of the space) or to the first person camera. The rule set doesn't just include code attached to virtual objects, but the code attached to the interface between user and system as well. On the right side of the interface, there is a list of code attached

to the first person camera. The perception of the user can be changed through different parameters: walking speed, height, field of vision, and so on. On the bottom left is the structure of triggers, which listen for certain events, as well as describe how code will be applied when such events happen. This means that the system can react to the user's behavior and the properties of the environment, objects and even characteristics of the first person perspective can change based on user input. For example, a chair might be brown and hard. But a chair in the rule set might be hard, brown, and play music when a user looks at it. Or it could explode if a user looks at it. Or it could move farther away from the user if she approaches it. The rule set of such objects does not need to resemble the rule set of physics or reality. Recall Arnheim's approach to realism: that 'realism' is not real, but the maker of a work has configured different elements so that it appears real. In other words, our realistic expectations could be met, but even if they are they have been constructed to do so. Realism requires just as much coding and rule construction as any other behavior. This new image is not only a data set. It is also a responsive rule set. The image as rule set has three aspects; *processual*, *operative*, and *embodied*.
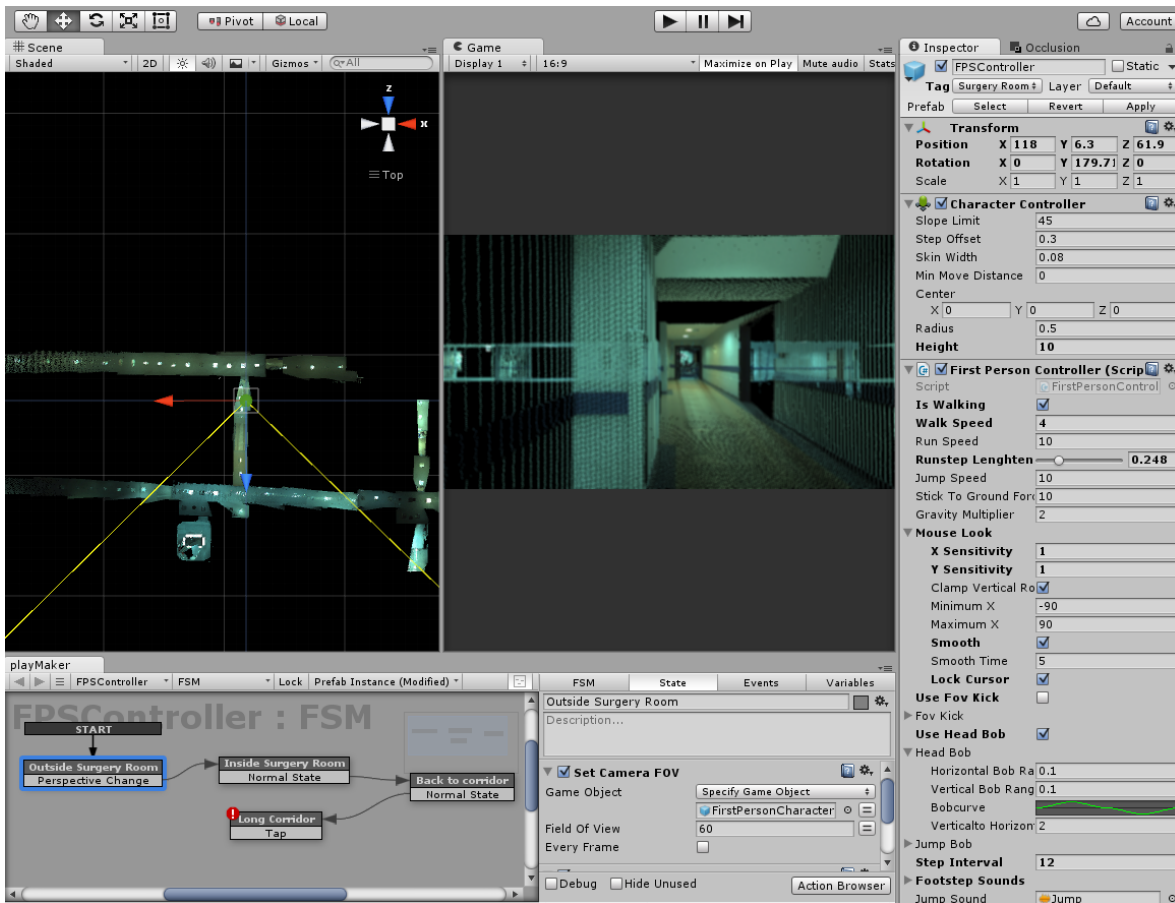
Figure 2.13: Unity Interface of *Hospital with one entrance*

2.4.1. Processual

The explorable point cloud visualization in *Hospital with one entrance* is what Schroeter

calls an "interactive transplane image" as it can respond to viewer input (377). The images are

constantly reprocessed and recreated: unlike a film or photograph, a new image is generated in

response to each modification of the viewer's position, giving the illusion of responding as a

physical reality would. The interactive transplane images "present virtual space … in a geometrical-

optical way … yet they also enable navigation … through this space … via their performative

actualization" (385). The tension between spatial and planar is apparent in the fact that such images

only create a virtual space through navigability.[23] This in-betweenness turns the interactive

transplane images into threshold entities as they occupy a ground between an image and object.

However, they are only presented on a screen, and the spatiality is maintained through constant

"processual intermediation," by the computational system endlessly presenting a new image in

relation to user's movement (Uricchio 34). In this sense, these transplanar images create a link

between time and space through navigation. As a viewer moves through navigation or attention,

these visualizations respond, and are rendered to the viewer differently. Because of this processual

relationship, the environment is continuously created with the user's input.

2.4.2. Operative

One aspect of virtual images is that they are interactive: not only because they are created

processually, in response to the viewer's positioning, but they are also *operative*. The user can take

action and operate on the images. Hoelz suggests that these images are not just operated *on*, but that

they operate on us as well. These images cause us to act in certain ways. For example, I allow

Google Street View (an image), to *operate* my route through real space. At the same time, I

generate the image as well as Street View data.

> Farocki's concept of the operative image, coined in the context of his discussion of
>
> automated warheads, where the image functions as a guiding tool for target tracking and the
>
> real-time adjustment of a missile's trajectory, proves to be very *operative,* too, when placed
>
> within the field of locative media and digital surveillance. "Location services" such as
>
> Google Street View, enabled by real-time data processing and continuous exchanges
>
> between user location, GPS sensor, software, network and database, are based on the
>
> principle of the users' trajectories feeding back into the database. The result is what we

---

[23] The user can navigate through a mouse or through her body, depending on what type of hardware she is using.

could call, with reference to Paul Virilio (1989), a 'reverse operativity' which proves to be

the more problematic side of locative media applications: It is not only that *we* are operating

the world through Google's images, it is also and primarily that Google's images are

operating *us*. (Hoelzl)

Such digital images resonate with Schroeter's concept of the inverted panopticon of digital media:

it is not just the user looking at the images but the images looking back at the user. The virtual

space is an inverted panopticon. Unseen spaces are unfixed until a viewer turns to them. The user

can act on the environment.

2.4.3 Embodied

Hospital with one entrance is viewed through virtual reality head mounted display

technology (HMD). Virtual reality HMD is a display technology with three essential technical

attributes: 360-degree image space, stereoscopy, and head tracking (Sutherland 26). Stereoscopy

makes binocular vision possible, adding a sense of depth. The 360-degree image space offers a

panoramic viewing environment. Head tracking is the system's ability to respond to the head

movements of the viewer and change the image on the display accordingly. Schroeter claims that

the head-mounted display "connects the stereoscopic image with the 360-degree panorama. This

connection is historically new. It becomes possible only because the image changes with the

movement of the viewing subject's head" (384). With head tracking, the immersive capabilities of

stereoscopy and 360-degree image space are combined. Virtual reality HMD technology falls under

the category of what Crary calls "physiological optics"; it is optimized for the human observer and

"the idiosyncrasies of the 'normal' eye" such as "peripheral" and "binocular vision" (16). This

section is titled "embodied" because the images are closely linked to the viewer's embodied

perceptual system, foregrounding the subjectivity of the human body.

Virtual reality optimizes the technology for not only human vision, but also human movement. With positional tracking technologies, such as optical trackers or Google's Project Tango, a virtual reality system can enable images to react to the user's body movements and position. Then, the co-responsive images situate the user's body in a virtual environment. Furthermore, through peripherals, such as hand controllers, eye tracking systems, other bodily gestures can be taken as inputs to the system.

Virtual reality is a mixture of various technologies, but is still predominantly a visual medium. There are various bodily means of input in a virtual reality experience: head/body movement, gaze, hand gestures, to name the most popular ones. However, the output in a virtual reality system is primarily audio-visual. It is an image that reacts to multiple inputs, multiple parts of the human body, such as head, body, eye, hand controller. Through reverse operativity (the system's perception of the user), the user's body and gaze is constantly tracked in a virtual world. A space can react to an observer, and to the observer's attention. The images and the sound are able to reorient the viewer's internal representation of her body in a virtual reality experience by visually responding to bodily input. With the virtual reality hardware, the environment is created continuously with the user's bodily input and the user can act on the environment with her body.

I call the embodied experience of softimages while using virtual reality systems *machine vision perspective.*

2.5. Machine Vision Perspective

The report on laser scanners emphasizes that these instruments capture "evidence", evidence that is missed by the human eye but caught by the machine. This can be seen as the continuation of claim for mechanical objectivity. However, it is also reminiscent of what Paul Virilio calls "sightless vision": images created by machine viewpoints (59). Jonathan Crary also notes that such

imaging technologies relocate vision "to a plane severed from human observer" (1). *These technologies produce a wide array of non-fiction imagery that complicates the relation between the observer and representation.*

Schroeter suggests some ramifications of how machine vision represents the world. He quotes Lacan writing about how the outside gaze can determine the conception of the self: "What determines me, at the most profound level, in the visible, is the gaze that is outside" (142). It is not a surprise then that a recent art piece, created with a Kinect scanner, was centered on the audience seeing the image of herself as a transplanar image. In 2010, artist Kyle Mcdonald created the *The Janus Machine*. The piece is a live-feedback photo booth, in which the viewer sits in front of a Kinect and sees her transplanar visualization on the screen. The piece allows the audience to watch her image captured from all sides as a spatial formalization; a case in which "one becomes acutely aware of the spatiality of one's own body" (Schroeter 131). This is a curious case, in which "people become objects of the machine gaze…" (130). Media theorist Mark B.N. Hansen writes that these images are non-perceptual, that they offer a "presentation of worldly intensity … independently of any act of perception by any being, human ones included" (Hansen). Seeing oneself through the machine vision prompts the viewer to imagine a different conceptualization of her body.

Virilio writes that machine vision should be considered "in relation to the philosophical question of the *splitting of viewpoint*, the sharing of perception of the environment between the animate (the living subject) and the inanimate (the object, the seeing machine)" (59-60). [24] There is something novel in combining 3D capture techniques with virtual reality HMD. Through the combination of the point clouds in *Hospital with one entrance* and the virtual reality HMD, the human observer is granted an inside look at the synthetic perception of the machine. Through the physiognomic optics of virtual reality HMD, the machine vision is embodied by a human observer.

---

[24] 3D capture is also a representation of invisible processes of a mixed reality future.

Seeing through a representation of machine vision prompts the viewer to imagine a different experience of her body: What it might be like to see the world as a machine sees it. Machine vision perspective is the ability to experience a machine's sight through a virtual reality system. This makes *embodied montage* possible, and is the focus of the next chapter. Embodied viewing of machine vision gives the animate human body to perceive through a radically different perspective: the perspective of the inanimate.

The viewer encounters the physical world mediated through computer vision, as in 3D laser scanning. This encounter is not unusual to many individuals who regularly experience it through computer screens, especially for professionals using 3D capture technologies, or amateurs experimenting with depth capture such as Kinect or similar technologies. However, it is more unusual to encounter the 3D representation of the world, the machinic vision of the real, through an embodied technology such as virtual reality. These images and the space, when combined with a virtual reality headset and optical tracker, can react to an observer's movement and attention. So what creative possibility does this ability of embedding code into objects, spaces and the first person camera give us? The synthetic view of machine vision experienced through an embodied display can provide a ground to expand and experiment upon conceptual models. This pairing allows for certain forms of interactivity to surface which can produce emotional and intellectual states. Manipulating these interactions can produce new types of experience that can lead to embodied abstractions not possible with the human body situated in the physical world, emerging models of speculative thoughts and new meanings. In the following chapter, I argue that *embodied montage* is the deliberate decoupling and recoupling of action and perception, in order to manipulate and use these pairings as a creative technique that can generate new meaning, as montage did for film.

Chapter 3:

Embodied Montage

"What is a symbol? It is to say one thing and mean another. Why not say it right out? For the simple reason that certain phenomena tend to dissolve when we approach them without ceremony."

(E. Wind)

(Opening epigraph of Raul Ruiz's *Poetics of Cinema*)

3.1. Introduction

In this chapter, I propose a framework for creating and analyzing works of virtual reality: *embodied montage.* This framework builds on the idea of montage from film theory. Montage is a technique for editing film clips together in such a way that new meaning can be created. Embodied montage is a renegotiation of action and perception that is possible within virtual reality systems; it produces new meanings by creating unexpected links between action and perception. This technique breaks how one perceives herself, and through that paves the way to creating new meanings. Supported by research in embodied cognition, the process by which sensorimotor, perceptual activities exist in a feedback loop with concepts and cognition, embodied montage suggests a modality for creating new meaning using virtual reality systems.

In this chapter, I will discuss the historical genesis and development of montage theory, propose Jean Luc Godard's *Goodbye to Language* as a transitional example where montage becomes physical, and then discuss various examples of embodied montage within virtual reality work. Embodied montage takes its name from film history, particularly Sergei Eisenstein's montage theory. *Goodbye to Language* breaks montage out of cinematic boundaries, and leads to my conceptualization of embodied montage. Machine vision perspective, a prerequisite for embodied montage, offers a representation of the environment as a malleable space that can be controlled by the body, and makes embodied montage possible.

3.2. Montage Theory

In the 1920s, Soviet filmmaker Sergei Eisenstein developed an approach to film editing called "montage." Rather than creating sequences in which shots follow each other seamlessly, he placed contradicting shots directly after one another in order to create a dialectical meaning. These two shots when combined create a third meaning that does not exist in either of the shots alone, the

whole ends up being greater than the sum of each shot. The heart of montage lies in creative juxtaposition, in counterpoint; the basic principle is to give "birth to concepts, to emotions, by juxtaposing two disparate events" (Eisenstein 58).

Eisenstein is in dialogue with the Kuleshov experiments of the early 1920s. Soviet formalist filmmaker and theorist Lev Kuleshov conducted film editing experiments with an audience. In one experiment Kuleshov edited the actor's expressionless face to be followed directly by three different shots: a woman, a soup and a coffin. The first shot of the actor is followed by the second shot and then the actor's face is shown again. Even though the shot of the actor's face is the same in every shot, the audience associates different meanings with each sequence: lustful, hungry, and sad when respectively juxtaposed with the woman, soup and coffin.

This experiment serves as a "proof that cinematic meaning is a function of the ordering of shots" (Prince and Hensley 59). The meaning is created not only through spatial composition but also through sequencing of shots; chronological composition. Sequential juxtaposition of images create a new meaning that does not exist in any single shot. Eisenstein took Kuleshov's experiment a step further and articulated one of the most significant theoretical and artistic theories in film editing, the theory of montage.

In "A Dialectic Approach to Film Form," Eisenstein writes: "To determine the nature of montage is to solve the specific problem of cinema. The earliest conscious film-makers, and our first film theoreticians, regarded montage as a means of description by placing single shots one after the other like building-blocks" (48). By the time Eisenstein was writing in 1929, continuity editing was already highly developed in the United States, especially by D.W. Griffith[25]. Continuity editing provides the rules for the most naturalistic ordering of shots; such as matching eye lines –

---

[25] Griffith and his film *Birth of a Nation* is an important reminder that formal techniques are not always used to expand people's minds in positive ways, but can also be used for manipulation, such as in this case, for odiously racist propaganda.

cutting in between actors as if they are looking at each other–, or the 180-degree rule which gives a more naturalistic feel of the space. Continuity editing "reinforces spatial orientation" and "represent[s] space for the sake of the story" (Bordwell 56). The visual grammar of continuity editing allows the audience to forget the artifice of filmmaking and be embedded into the narrative of the film (Dancyger 5). Recall the discussion of "immediacy" and "presence" from Chapter 1. Continuity editing is essentially mimetic. The editing style tries to use the cinematic experience to imitate the attention of a human. Continuity editing maintains a familiar space-time continuum, and presents the space and actions closer to a real life experience.

In his work and his theory of montage, Eisenstein breaks this notion of film editing by emphasizing dialectics and the importance of conflict in meaning creation. Rather than creating sequences in which shots follow each other seamlessly, he places contradicting shots directly after one another in order to create a new meaning. For Eisenstein "montage is an idea that arises from the collision of independent shots-shots even opposite to one another" (49) to "forge accurate intellectual concepts from the dynamic clash" between shots (46). Setting two images off in reaction to each other is the operating principle of montage. Montage breaks with mimetic representation and the traditional space-time continuum. Montage juxtaposes images in a way that isn't possible in everyday visual experience. It is a method to push the expressive capabilities of cinema as an art form.

For instance, in *Strike,* Eisenstein cross cuts the killing of workers with the butchering of a bull. The intercutting between these two images creates an associative link of butchering in both instances. The function of the butchering scene is to give meaning to the first, the killing of the workers. This also leads to a heightened emotional reaction in the audience.

Figures 3.1 & 3.2: Montage sequence in *Strike*: Killing of the workers & butchering of a bull

In another instance from *Battleship Potemkin,* Eisenstein cuts from a priest's cross to a captain's sword in order to draw a parallel between them. Similar to the tsar's army, the clergy were "inherently hostile to the workers' movement" (Taylor 59). This sequential coupling communicates an association between the clergy and the tsarist regime.

Even though Eisenstein's project of creating stabilized meaning through images did not always achieve its goal[26], montage turned out to be very fruitful in creating cinematic meaning and poetic resonances that cannot be reduced to or articulated by language.[27] Montage has a large impact on the cinematic imaginary. Eisenstein's theory was very influential, especially in the way that it articulated a film grammar and shaped how the discipline conceived of itself.

---

[26] A later film of Eisenstein, *October,* is packed with instances of montage. However the audience was very confused; the film was highly criticized. Another Soviet director Pudovkin called the film "such a powerful failure" (Taylor 73). What becomes evident in *October* is that images do not equal to a similar type of communication which is carried through language. Communication through images do not necessarily create stabilized meanings.

[27] I am thinking of how Nathaniel Dorsky articulates the mysterious effect of montage through certain moments in film history: "And who can forget Jean-Luc Godard's cut in *Contempt* (1963) from the auto accident of the red sports car to the glistening blue sea as Michel Piccoli ascends the cliffs? So much emptiness, so much spirit, so much narrative purpose is encapsulated by that cut. Which brings to mind the last cut in *Voyage to Italy.* From the climactic catharsis of the embracing Ingrid Bergman and George Sanders we cut to a seemingly insignificant detail, a uniformed town official standing amid the passing crowd. How unexplainable. How poignant. How disarming" (47).

3.3. *Goodbye to Language* and the Embodiment of Montage

Figure 3.3: Separation of the eyes in *Goodbye to Language*

Almost a hundred years later, in Jean Luc Godard's stereoscopic 3D film *Goodbye to Language*, there is a moment that pushes the cinematic medium out of its own boundaries. In one scene, a woman, Ivitch, and a man, Davidson, sit together, looking at a book. Ivitch's husband comes to the scene and holds her by the arm and takes her away. In this moment, the two cameras which set up the stereoscopic image, accurately distanced from each other to recreate binocular human vision, split from that position. One camera stays on the man, while the other follows the woman and her husband as they walk out of the frame.

The 3D image separates into two distinct images, each is seen by one eye only. This simultaneous juxtaposition of images by projecting them to different eyes "takes montage to a logical extreme"; the scene becomes an "in-eye editing" (Marsh). The montage here is neither

temporal nor spatial. It is embodied, occurring within the physical perceptual system of the viewer, incomplete until viewed. It is created through projecting different images and ideas onto different eyes. It is a montage that addresses each eye differently, a type of perception that is rare or impossible in everyday experience. It is a genesis moment of what I will call embodied montage.

As discussed in the *Machine Vision Perspective* chapter, virtual reality hardware makes it possible to get multiple bodily inputs from the user: head movement, body position, hand and arm gestures. The real-time 3D engine makes it possible to produce real-time responsive outputs in the form of image and sound in accordance with these inputs. Thus, the responsive real-time images and sound react to the user's input; a head movement produces a certain response in the real-time experience. Most current virtual reality works using real-time responsiveness use it to create a seamless, mimetic experience. For instance if a user turns her head to the right, the image corresponds to this motion and changes accordingly. Or if a user walks within a virtual reality experience, the images change accordingly so that a feeling of walking in a 3D environment is created.

However, several artistic works try to break away from the seamless, mimetic experience that many virtual reality creators seek to create. These works expand virtual reality's language, and employ a technique that I will refer to as embodied montage. Specifically, they do this by creating new ties between user perception and actions in the virtual system.

3.4 Embodied Montage

Embodied montage is a technique that can be used in real-time virtual reality experiences. It is characterized by the decoupling and recoupling of action and perception using a virtual reality system, in order to generate new meanings. The unexpected links between action and perception

drive conflict and dialectic rooted in embodied experience.[28] In an instance of embodied montage the link between action and perception is intentional but does not mimic the real world experience.

The combination of softimages and virtual reality HMD can separate one eye from the other, separate height from the body, and separate the field of vision from sight. The rule set, that governs interaction with the data set, can reconfigure the body and its relation to the environment. Thus, an action, even an ambient action like *standing*, could result in a completely different perceptual experience. In combination, softimages, HMD, optical trackers and hand controllers can separate movement from vision. Such separations allow the artists to make novel combinations between these characteristics and actions of embodiment and their consequences.

Embodied montage is important in two ways: it can create new conceptual structures and meanings driven by embodied cognition; and it offers a response to the narrative paradox. First, embodied cognition is the notion that our lived experience in the world, through our bodies, sensorimotor abilities, and perception, plays a crucial role in structuring our conceptual life. Meaning is created through "patterns of embodied experience and pre-conceptual structures of our sensibility (i.e., our mode of perception, or orienting ourselves, and of interacting with other objects, events, or persons)" (Johnson 14). Embodied cognition argues that we think not just with our minds, but also with our bodies. According to embodied cognition, we develop abstract thought through pre-conceptual cognitive structures that "emerge from the recurrent sensorimotor patterns" (Varela 173) called image schemas: "An image schema is a recurring, dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience" (Johnson xiv). Cognitive scientists Mark Johnson and George Lakoff developed this theory of

---

[28] Embodied montage is reminiscent of and indebted to philosopher Gilles Deleuze's concept, the time-image. In post-war cinema, the action and consequence logic of classical cinema is broken. Shots do not follow each other in such logic; a sonic and optical system that loosens the narrative emerges. Time-image also decouples action and perception—yet does not make new linkages.

embodied cognition through looking at how humans construct metaphors. Abstract concepts and metaphorical thinking arise from image schemas, patterns of sensory-motor experience. For example, "up" is commonly associated with "happy," and "down" with sad. That is related to the human body and its vertical posture. The reasoning for this analogy for Lakoff and Johnson is that "dropping posture typically goes along with sadness and depression, erect posture with a positive emotional state" (15). We use metaphors to generate new meaning, which comes from human embodiment and lived experience. How we think, and who we are, depends on our bodies and their relation to the environment. The promise of embodied montage lies in the idea that *conceptual structures, how we think and thus who we are, can be altered by changing the relationship between action and perception, thus the body and the environment.*

Second, the trade off between interactivity and narrative experience is documented in the literature on interactive narrative as the *narrative paradox* (Aylett). A world can be completely interactive, but this makes it very difficult to control the pacing, timing, and content of a narrative. Inversely, a very tightly scripted experience approaches the sensibilities of a film, losing interactivity. However, expressive computational techniques suggest that there is a third path. In *Phantasmal Media,* Fox Harrell coins the term *agency play* as an expressive computational technique. Through agency play "a user's power to take meaningful actions is mediated through structures provided by the computational system" (269). One layer of agency play is the *agency relationship,* which is the operational relationship between "user actions and system actions" (270). This approach helps to negotiate how a work may both create meaning, and be interactive. In the preceding chapter, I defined machine vision perspective as the ability to experience a machine's sight through an embodied viewing system. This organization, where a user *sees* herself and the virtual environment as processed by a computer, allows for creative intervention into action and

perception, where there is an agency relationship between the user's body and the computational environment. I am referring to this approach as embodied montage.

Embodied montage is reminiscent of digital synaesthesia[29] or sensory substitution[30], but distinct from these two ideas. It does not attempt to translate one sense to another, such as making one see red by giving her tactile stimulation (sensory substitution), or translating music into visual animation (digital synaesthesia). Instead, embodied montage creates new unexpected links that exist in a dialectic way. For instance, a tactile stimulation can be linked with the experience of seeing red. However one does not cause the other; instead, they compliment each other. Furthermore, embodied montage is in close dialogue but also separate from what Sutherland calls *intentional looking*, an approach in virtual reality that creatively modifies the act of looking (Sutherland 51). Some instances of *intentional looking* are also instances of embodied montage; however, *intentional looking* is specific to the use of gaze: both from the user towards the environment, and from the environment towards the user, and includes other elements of composition (lighting, framing) aside from the computational aspects. Embodied montage is a more specific approach, in that it deals only with the new linkages between action and perception, specifically made possible by real-time computation, and a broader approach, in that it can be applied to other modalities besides looking that are also employed in virtual reality.

I will elaborate on embodied montage through several examples.

---

[29] Digital artworks through the digital code are endued with a mode of translating sensory dimensions directly from one sense to another, e.g. the translation of digital sound into digital imagery via the binary code ("What is digital synesthesia?")

[30] Sensory substitution means a transformation of the characteristics of one sensory modality into stimuli of another sensory modality. It is hoped that sensory substitution systems can help people by restoring their ability to perceive certain defective sensory modality by using sensory information from a functioning sensory modality.

In Oscar Raby's *Assent*[31], the user takes the point of view of the artist's father, a former soldier who served in the Chilean army during the Pinochet military regime. The piece recreates the experience of witnessing a mass execution. In order to navigate through space, the user looks at a particular place for several seconds, and then moves towards it, a linkage made possible by real-time computation.[32] This interaction mechanism combines the action of looking with the consequence of moving (perceiving movement), creating a non-mimetic link between these two acts. This moment links looking and moving, giving new meaning to act of looking; the montage creates a third meaning which is not inherent in neither the action of looking and consequence of moving. What emerges in this juxtaposition is that looking means moving thus acting and being involved: similar to the artist's father who was involved in the mass execution by witnessing it. By creating a non-mimetic meaning for *looking*, Raby is utilizing a new form of montage: embodied montage.

---

[31] Created by Oscar Raby in 2013, *Assent* is one of the most critically acclaimed real-time virtual reality documentaries.
[32] This is also an example of intentional looking. However, embodied montage emphasizes that a third meaning is created, not inherent either in looking nor in movement.

Figure 3.4: User moves to where she looks in *Assent*

Marshmallow Laser Feast's *In the Eyes of the Animal*[33] allows the user to adopt the perception of an insect by using embodied montage. The user is put in a 360-degree environment of a point-cloud forest, acquired by laser scanning. The environment appears as large abstract points at first. As the user stares at the points, the points under her gaze become smaller and create a perceptible environment. In the case of *In the Eyes of the Animal*, the act of staring is combined with the consequence of making the environment gradually visible. This pairing creates the effect of eyes slowly adapting to the environment, something that is not inherent in human vision, and thus speculates on how insects might perceive their environment.

---

[33] Created by the creative collective Marshmallow Laser Feast in 2015, In the Eyes of the Animal is an experimental virtual reality documentary. Some iterations of the piece experiment with headset and site-specific installation design.
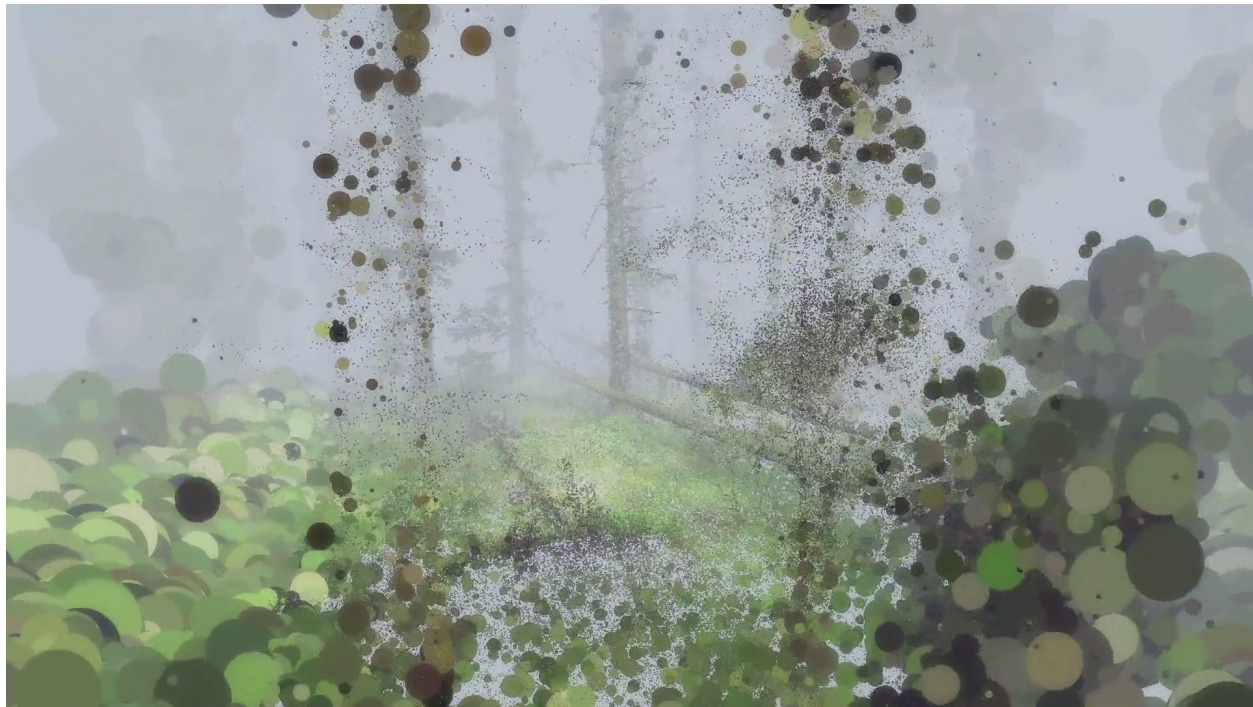
Figure 3.5: Objects become sharper as the user stares at them in *In the Eyes of the Animal*

Similarly *Notes on Blindness*[34] attempts to recreate the experience of a blind person through embodied montage. In the piece, as the user stares at dark and empty places in the environment where positional audio can be heard, the shapes of objects begin to take form. In this piece, the act of staring is combined with revealing unseen shapes. This pairing recreates a blind person's experience of perceiving objects by attending to the sounds they create.

---

[34] Created by Arnaud Colinart, Amaury La Burthe, Peter Middleton, and James Spinney in 2016 as a companion piece to the film with the same name, *Notes on Blindness* has been a festival favorite, winning the Storyscapes Awards at Tribeca Film Festival 2016.
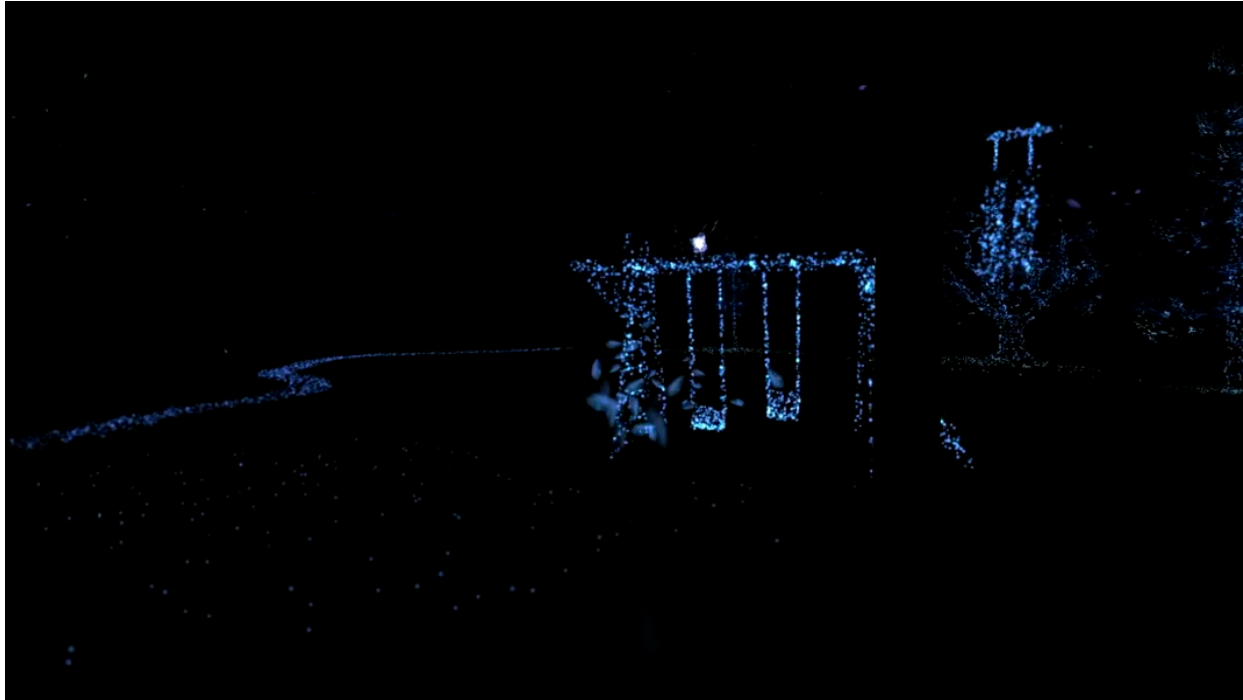
Figure 3.6: Objects form as the user looks towards the sound in *Notes on Blindness*

In Specular's *Blackout*[35]*,* the user finds herself in a stuck NYC metro car with a number of people. As the user stares at different people, she starts to hear their thoughts. Some of them are complaining about the subway while others think about their daily plans. In this instance, staring is connected with hearing another's inner thoughts. This combination creates a psychic experience, the ability to know someone else's train of thought by looking.

---

[35] *Blackout* is a project by Specular Studio: James George, Mei-Ling, and Alexander Porter. As of writing of this thesis, it is still in development. Thanks to Alexander Porter, I had the opportunity to a experience a work in progress scene.

Figure 3.7: User can hear thoughts of other people by staring at them in *Blackout*

The examples just discussed primarily use the direction of gaze to produce novel effects, but embodied montage can also be found through other actions. There are limited examples available, but Rachel Rossin's *Timescrubbing* and several techniques employed in *Hospital with one entrance* point towards what is possible.

In *Timescrubbing*[36]*,* in addition to a virtual reality HMD, Rossin also utilizes the optical trackers in the HTC Vive system to track user's movement in space. The environment is a 3D rendering of a living room; the everyday objects in the house are presented as imperfectly rendered 3D objects. The imperfections add to the quotidian nature of the environment. The user's movement in space not only controls bodily movement but it also controls passage of time. The body becomes an interface to control temporality. If the user does not change places, time does not

---

[36] *Timescrubbing* is a work in progress by Rachel Rossin. I had the fortune of experiencing an early prototype of the piece in Rossin's studio.

pass. For example, unlike a video, where time is fixed and linear, time itself could be controlled by a user's movement within a space. In games such as *A Slower Speed of Light*[37] and *Braid*[38] control of time functions as a kind of "game mechanic." In virtual reality it can be more tightly linked to embodiment and perception—not necessarily a control intellectually deployed.[39] In virtual reality the control of time can be another capability or affordance of the body. This pairing between movement and the passage of time creates a completely novel approach to the quotidian in *Timescrubbing*. Combined with the 3D objects, this instance of embodied montage offers a new aesthetics through which to address boredom and habitual space. Furthermore, this instance is also a moment of merging between the physical world with the conceptual world. Mark Johnson writes that humans think about time through the analogy of moving in space: "We (adults) conceptualize time via deep, systematic spatial movement metaphors in which the passage of time is understood as relative motion in space" (28). The embodied montage in *Timescrubbing* allows conceptual abstract thought to become physical.

In *Hospital with one entrance*, I attempt to use embodied montage in a fashion that includes movement. In the piece, when the user enters the operation room, the field of view slowly expands. This instance links moving into a specific space with a change in vision. The effect is a mix of subtle dizziness and a relief, a lessening of claustrophobic space. This pairing aims to reveal an affective state about the process of waking up from anesthesia; a feeling between disorientation and relief.

---

[37] *A Slower Speed of Light* is a game developed by MIT Game Lab in 2012, demonstrating the effects of special relativity.

[38] *Braid* is a platform game developed by Jonathan Blow and released in 2008. It is considered as one of the most innovative and popular indie games.

[39] A similar mechanism plays a role in Bruno Nadeau & Jason Lewis's piece *Still Standing.* Even though *Still Standing* is not a VR experience, it receives input from participant's body through motion capture. In *Still Standing* stillness acts as an interaction mechanism—the user can increase the legibility of the video projected text by standing still. http://collection.eliterature.org/2/works/nadeau_stillstanding.html

Another technique in *Hospital with one entrance* is influenced by *Goodbye to Language.* If the user spends too much time in the operation room, the image on one eye becomes blurred every three seconds for one fifth of a second and then goes back to initial, non-blurred stage right after. Because of binocular rivalry, the brain cannot register the blurred image, and the user still sees the operation room clearly. However, the blurring of the image on one eye creates a haptic effect, a feeling as if someone is tapping on the user's eye. This creates a mild discomfort that encourages the user to leave the operation room. In this instance, a characteristic of embodiment, binocular vision is disrupted by projecting different images to each eye in order to create perceptual discomfort and a simulated haptic feeling.
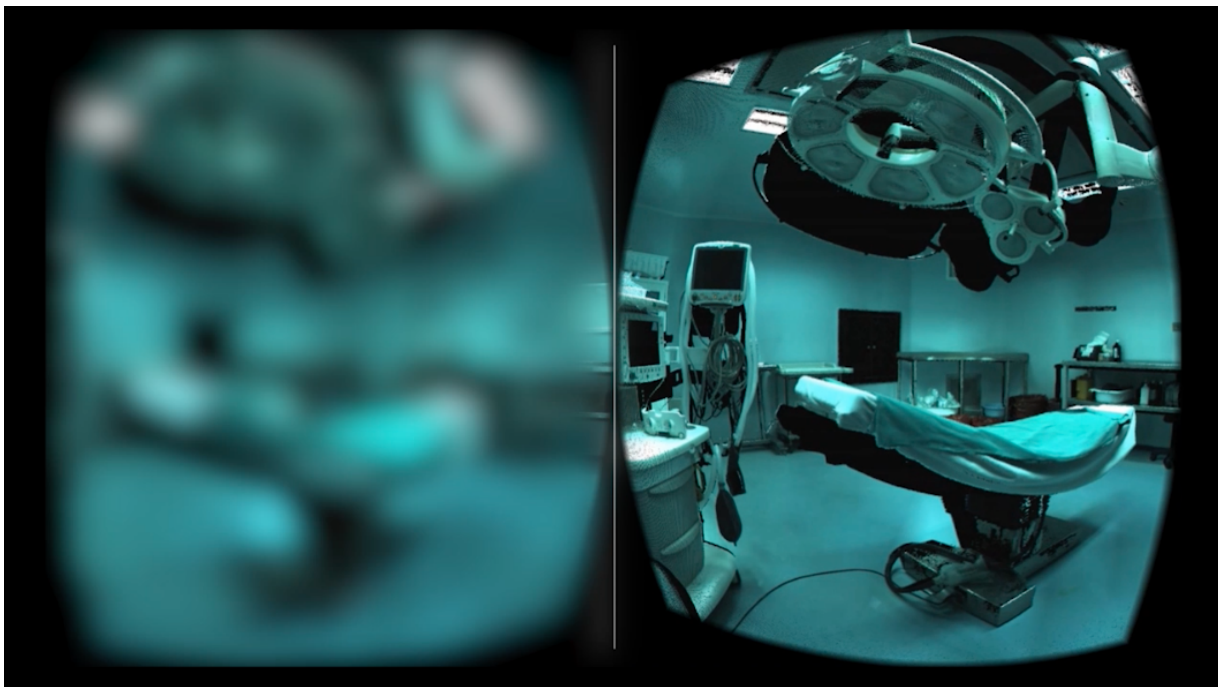

Figure 3.8: The blurring of the image on one eye in *Hospital with one entrance*

These are rather early examples but this principal has broad applications: by modifying the rule set of the system any movement or interaction can be abstracted, and be caused to trigger something unexpected, including body movement, gaze, and bodily gestures. Virtual reality

systems allow embodied montage to be physically experienced, but it is not the first time this notion appears in creative work. One instance from mythology is the story of Orpheus. Orpheus descends to Hades to take back his dead wife Eurydice. The god of the underworld grants his request on one condition: that as Orpheus leaves Hades, Eurydice will follow him, but he must not turn to look back at her. Towards the end of their journey, Orpheus cannot fight his curiosity and looks back at Eurydice to see whether she is following. Because of this, Eurydice stays in Hades and is gone forever. In this case, the action of looking is imbued with a special power: it is coupled with losing Eurydice. Similar examples where the action is tied to an impossible consequence are present in many other stories: In the Bible, Lot's wife looks back when leaving Sodom and turns into a pillar of salt; Medusa turns onlookers into stone by making eye contact; Midas turns everything he touches into gold. The association of some impossible ability with a mundane, embodied act is present in how we think of psychic abilities as well: such as a psychic looking at someone and hearing her thoughts; or in magical thinking: such as Werner Herzog's constant act of walking in an attempt to protect his friend from death as narrated in his 1978 book *Of Walking in Ice*; or quantum theory: in Heisenberg's uncertainty principle, observing is fundamentally linked with altering. These instances would provide useful strategies to tell stories through embodied montage.

Embodied montage offers a counter-part to whole-body interaction mechanics, which are similar to continuity editing. They provide tools to navigate in a seamless comfortable experience. Embodied Montage uses the same principle to expressive ends. The other aspect I want to talk about regarding embodied montage is creating pairings of action or perception with non-action or non-perception, pairing with blanks, which offers an expressive technique through disembodiment.

3.4.1. Pairings with Blanks: Disembodiment through Embodied Montage

Disembodiment through embodied montage is the "negative space" of the approach. In a widely circulated article, Oculus Story Studio articulates about "the Swayze Effect" which they think is a major obstacle to expressive storytelling in virtual reality: "The Swayze Effect … describes the sensation of having no tangible relationship with your surroundings despite feeling present in the world … Basically, it's the feeling of yelling "I'm here! I'm here!" when no one or nothing else around seems to acknowledge it" (Burdette).[40] This effect is caused when the user's actions have no consequences in the virtual world. For example, she can look at a character but the character does not respond, or she can walk in a virtual space but she can't use other bodily movements, such as hand gestures. I argue that this is not a problem that needs to be solved, but an expressive potential that should be explored.

Disembodiment can be seen as another variation on the *agency relationship* between the user and the system. In this case, the system takes agency over the user's body; the user's bodily agency gets limited and in some cases the bodily affordances become reduced. I will look at *Project Syria, The Enemy* and *Phantom* to articulate how this can be used to expressive ends.

Nonny de la Peña's *Project Syria*[41] recreates the experience of a bombing in Aleppo, Syria. *Project Syria* uses the virtual reality headset and optical tracking to positionally track the user. This system allows the user to walk within the virtual 3D environment. The user starts off in a 3D modelled street in Aleppo. After a brief chance to walk and explore the street, there is an explosion, and a cloud of smoke covers the street. This is a crucial moment of disembodiment within the piece. The explosion is created through sound and image only; no other bodily sense is stimulated. This is

---

[40] This effect is named after Patrick Swayze, whose character in the 1990 film *Ghost*, goes through a similar experience.
[41] *Project Syria* is created by Nonny de la Peña in 2014. This immersive journalism piece is commissioned by The World Economic Forum to raise awareness about the children displaced because of the Syrian civil war.

of course due to the nature of simulation. Simulations do not replicate reality point for point, but rather only selected elements of it. After all, complete replication would have the same consequences of reality, losing the learning advantage of an always partial simulation. Because of this a simulation is essentially about "the disappearance of reality" (Schroeter 381). In *Project Syria* when the rocket strikes the street, the body of the user is unaffected; the action of explosion is detached from the perception of feeling pain. Even though debris strikes the virtual body of the user, the physical body is not affected. This is a ripe creative moment that virtual reality pieces should acknowledge and make creative use of. In *Project Syria*, such disembodiment could signify the impossibility of fully understanding the situation, emphasizing the "grave human rights violations taking place" (Tortum).

Figure 3.9: Bomb explosion has no effect on the user's body in *Project Syria*

Karim Ben Khelifa's piece *The Enemy*[42] provides another instance of disembodiment. This

piece is one of the most major productions in non-fiction virtual reality so far. The piece is still in

development as of writing, and elements of the installation have changed several times. Here, I am

referring to the version of *The Enemy* that was exhibited in Tribeca Storyscapes 2015. In *The*

Enemy the user finds herself in between a Palestinian combatant and an Israeli soldier. Unlike

Project Syria, The Enemy does not attempt to bring the user into the setting of the conflict,

rightfully concerned with the issues of recreating a potentially traumatizing event. Rather it puts the

user in a rather abstracted room with two combatants on each side. Through optical trackers the

user can move within the virtual room in between the fighters. When she approaches one of them,

an interview starts playing and the user listens to their recounting of the conflict. The eyes of the

virtual characters follow the user's movements, lean back when a user gets too close and look

straight into the eyes of the user. This aspect adds a haunting presence to the piece. The boundaries

of the virtual room are determined by the location of the motion trackers; the user cannot walk

through the walls because that is where the virtual room ends. However, the user can walk into the

fighters. Even though most users do not do this and respect the distance between themselves and the

virtual characters, I think this is a curious and important element in *The Enemy*. If the user walks

into one of the fighters, the whole virtual world turns red and the field of view changes. This

moment is the only indication that the virtual body is different than the physical body. The virtual

body of the user is disembodied in the sense that she can walk into people. This moment breaks the

feeling of immersion but it also serves as a poetic moment, similarly signifying to an impossibility

present in a simulated world: to really be with someone in the same place, to look through the eyes

of someone else.

---

[42] *The Enemy* is a virtual reality and augmented reality project directed by photojournalist Karim Ben Khelifa. It is still
in development, however, several iterations of the project have been publicly exhibited in 2015 and 2016.

Figure 3.10: The user can move through a virtual human in *The Enemy*

Both pieces offer agency play between the body in physical world and the body in the virtual world. In *The Address of the Eye*, film theorist Vivian Sobchack discusses film experience from a phenomenological standpoint. Delving into the bodily experience of the film spectator, she elaborates on the "double occupancy of cinematic space" (10). In a movie theater, the spectator engages with the visible world presented in the film as its own objective reality as well as her own manner of watching the film; this leads to a dialogical relation of spectator viewing the film as her own vision as well as of another's. The viewer occupies her own body and the body of the camera at the same time. Sobchack writes: "each film projects and makes uniquely visible not only the objective world but the very structure and process of subjective, embodied vision" (298). Her work on the bodily perception of film experience resonates very closely with the disembodiment in virtual reality (Boyacioglu). The double occupancy in virtual reality is not merely a question of

vision. It is more somatic, as the viewer's body exists both within the virtual world and the physical world. The immersive visual field and responsive images create a sense of presence within the virtual world whereas non-audiovisual modalities (like tactile sensations) present a different presence that is still rooted in the physical world. The virtual reality experience offers a "double presence," which is pushed further in *Phantom.*

Phantom (kingdom of all the animals and all the beasts is my name)[43] is a virtual reality installation by artist Daniel Steegmann Mangrane, placed in a small rectangular room with gallery-white walls. On the floor, there are two concentric circles drawn with gaffer tape. A blue cord hangs down through a small hole in the ceiling. Halfway to the ceiling, on each corner of the wall, motion sensors emanate a faint blue light. The cord carries a headset, which hangs in the center of the circle. When in use, a viewer stands in the center of the circles wearing the headset. The cord moves slowly as she turns her head. As she takes larger steps; she crosses the first circle. The cord tightens and as she gets to the edge of the larger circle, the headset begins holding her back. The limits of her range are marked by the outer circle of tape. Outside of the installation, looking in, the participant appears to be part of the piece, moving slowly like a sleepwalker.

When the user puts on the headset the vertical linear formation of the cord of the head mounted display and the viewer's body become the tree of a trunk within the virtual reality experience. The user finds herself in a trunk of a point cloud representation of a forest. When she looks down to see the boundary of the trunk, she realizes that it corresponds to the circle drawn on the floor of the installation. The physical world is formally converted into the virtual world.

However, the virtual body has reduced affordances. Nothing reacts to the body as a material: the user can walk through the bushes and trees that are made from point clouds and she

---

[43] *Phantom* is a 2015 virtual reality installation by Daniel Steegmann Mangrane. It is not widely exhibited or publicly available at the time of writing. I had the opportunity to experience it in the 2015 New Museum Triennial.

cannot operatively interact with anything in the virtual world. And yet, the environment does processually respond to the viewer and allows her to navigate the environment, changing her perspective and position within the space. In *Phantom,* the body becomes the negative space in the forest. This relation is a different way of relating to images; it is a reconsideration of the body, in which the movements of the body and the bodily presence is felt only through the virtual objects in the environment. The body is marked by an absence; when the viewer walks away from the initial position, there is a black hole. The black hole in the center of the virtual environment becomes the sight that explains *Phantom. Phantom* is a contemplation of negations: in the lack of human visual perception, the lack of bodily integrity, the lack of Newtonian physics and Euclidean space. The double presence of the viewer is suspended between the virtual and the physical. *Phantom* offers the possibility for contemplation for a holistic non-anthropocentric outlook, for a world where sensation and subjectivity doesn't end at the physical limits of the self. A world in which the body and the environment blend into each other, and the body of human observer is not fixed.

Figure 3.11: The black hole at the initial position of the body; at the original point-cloud capturing

position of the laser scanner in *Phantom*

3.4.2. Categorization of Embodied Montage

The following categorization of additive and subtractive pairings in embodied montage is by

no means definitive; rather it is presented here to give a clearer understanding of different valences

of the technique and open up discussion for further creative explorations. Embodied montage can

be additive or subtractive: it can either create a new pairing between action and perception within

the body or between the body and the environment, or it can break an already existing pairing.

Embodied montage can happen within the body: it can combine bodily modalities such as staring

and moving (*Assent)* or staring and hearing (*Blackout)*. [44] It can happen between the body and the

environment. The body can act on environmental perception, creating a new pairing, such as the act

---

[44] I haven't seen any intentional creative use of subtractive embodied montage within the body. Most of the virtual reality pieces do this unintentionally. For instance, in all the works analyzed here, the user cannot use her hands: this already breaks an already existing pairing within the body.

of staring causing the illumination of objects (*Notes on Blindness)* or movement of the body causing a change in time (*Timescrubbing).* Alternatively, a preexisting relationship between the body and the environment can be effaced; a user can move through the trees (*Phantom)* or a virtual human (*The Enemy*) breaking the rule of physics of the real world. Also the environment can act on bodily perception. Entering a room can trigger a change in field of vision and create an additive montage (*Hospital with one entrance)* or a bomb explosion can have no effect on the user's body and create a subtractive montage (*Project Syria).*

Embodied Montage			
	Body acts on bodily perception	**Body acts on environmental perception**	**Environment acts on bodily perception**
Additive	*Assent*: Staring causes locomotion *Blackout*: Staring causes hearing	*In the Eyes of the Animal*: Staring makes the image sharper *Notes on Blindness*: Staring causes illumination of objects *Timescrubbing*: Moving causes change in time	*Hospital with one entrance*: Entering certain rooms causes a change in the field of view. Extended presence in a room causes "eye tapping"
Subtractive		*Phantom*: The trees and bushes are insubstantial, a user can move through them *The Enemy*: The user can move through a model of a human	*Project Syria:* Bomb explosion has no effect on user

Figure 3.12: Categorization of embodied montage

3.4.3. The Promise of Embodied Montage

Just as machine vision perspective offers a different conception of the self, embodied montage offers a different experience of the self. The images of virtual reality are not fixed. The body in virtual reality is also not fixed.

Embodied Montage then leads to several outcomes:

1) Embodied montage is an expressive technique for virtual reality. Similarly to montage in film, it allows creators to construct narrative meanings by juxtaposing unexpected actions and perceptions. With the use of embodied montage, a virtual reality experience can be interactive but also structurally convey a narrative. It is one way of overcoming the narrative paradox.

2) It is possible to create new meanings by creating new links in the sensory-motor schema, by coupling action and perception in an unexpected way. With embodied montage novel relations within the body and between the body and the environment can be created. The body can be renegotiated and this would alter the *image schemas,* the preconceptual cognitive sensorimotor structures affecting how we think. This method could potentially offer a contemplative alternative to the body-centric and human-centric world-view. In short, embodied montage offers a new way of thinking of human body as an interface. It reorients the body for expressive means.

3) Embodied montage also offers a critical distance to the process of mediation and becomes a tool to interpret the experience. A non-fiction virtual space should make its system, the operations behind its system and the capture mechanisms visible. This

can occur through machine vision perspective and embodied montage, through glitchy, imperfect scans, in an attempt to show the apparatus, through creating novel pairings between action and perception and let the user interpret her body. It is important to do this because it enables viewers to enhance their literacy. Conventions in virtual reality are currently aimed at naturalistic realism. But virtual reality also needs more emphasis on the reflexive mode in order for us to have a critical distance to the mediation. Since these are, to a degree, alienating techniques, they put a user at a critical distance to the mediation. In this way, the user has the ability to contemplate the mediation and preserve the distance between representation and real. It does not recreate reality but provides formal elements to interpret the experience. Contrary to immediacy, the medium is made apparent. Embodied montage breaks with realism to create new, unexpected meaning through juxtaposition and conflict.

Conclusion

"I love the idea of technology and culture moving faster than the understanding of those mediums by people."

(Ryan Trecartin)

In his talk at the *Virtually There* Conference at MIT, architect Marcos Novak made a resonant analogy between virtual reality and the telescope: "If you have a telescope and you point it to your shoes, you are going to see a blurry pair of shoes. You have to point it to the stars to make it useful." Implicit in this claim is the idea that we are not using the virtual reality technology properly, although it has an immense potential for discovery. This notion is prevalent among other important figures in the art and technology world. For instance, when asked about virtual reality, filmmaker Werner Herzog replied: "In this case, we do have a technology, but we don't have any clear idea how to fill it with content" (House). Virtual reality is capturing the attention of many different fields: entertainment, journalism, art, and documentary to name a few. Yet there is no agreement on how virtual reality can be effective and lead to new territories of human perception and communication. I hope this thesis offers a useful framework to approach virtual reality. Drawing on film theory, media history and cognitive science, I have argued for the framework of *embodied montage*. This framework creates a territory for creative expression that challenges embodied cognitive structures and could effectively use the medium in ways that are distinguished from other media forms.

Chapter 1 argues that the goal of *immediacy* in virtual reality does not implicitly provoke expressive use of the medium. The writings of Walter Benjamin and Jonathan Crary as well as the scientific discussion of the 19th century show that *new media technologies can lead us to create new forms of knowledge,* by breaking accepted rules about how we perceive the world. Immediacy, and forms of media creation which attempt to efface the material reality of the medium, obscure this practice. Virtual reality's "realism" is a construct, and through acknowledging this, the medium can be pushed to more experimental and novel ends. In order to diversify how we think about the medium, I attempt to dissociate the technology through analyzing images. My methodology

involves a critical making project *Hospital with one entrance* and theoretical approach following media theorists Hoelzl & Marie, Uricchio, Sutherland, Schroeter, Grau, Elsaesser, Manovich and Virilio.

Chapter 2 starts by describing the experience of the non-fiction virtual reality work *Hospital with one entrance.* In order to analyze the piece, I adapt the framework of "softimage" proposed by Hoelzl & Marie, and approach the images of the piece in two main categories: image as data set and image as rule set. Image as data set offers four affordances; the images are transplanar, they are between planar images and spatial objects; referential, they have a real-world, indexical referent; virtual, they are captured through non-optical means (at least in the sense of photography); and contingent, their perspective, transparency and spatial relations are subject to change. These factors lead to the elements of the image as rule set. The image is processual, continuously created with the user's input, operative, the user can act on the environment, and embodied, with the use of virtual reality hardware the user can act using her body. All these aspects combined generate the condition of *machine vision perspective,* the act of experiencing softimages through a virtual reality system. I argue that this provides a ground to experiment with embodied modes of thinking that are not possible with the human body situated in the real world. This is possible through *embodied montage*, the decoupling and recoupling of action and perception, in order to generate new meanings, similar to montage in film.

Chapter 3 proposes a framework for creating and analyzing virtual reality works: *embodied montage.* This framework builds on the technique of montage from film theory and is supported by research in embodied cognition. Embodied montage is a renegotiation of action and perception that is possible within virtual reality systems; it produces new meanings by creating unexpected links between action and perception. The chapter discussed the historical genesis of montage theory with

Eisenstein, proposed Jean Luc Godard's *Goodbye to Language* as a moment when montage becomes embodied, and then analyzed examples of embodied montage within virtual reality work. The promise of embodied montage is threefold. It can overcome the narrative paradox and serve as a technique to structurally convey a narrative meaning. It can create new links in the sensory-motor schema, between action and perception, and affect how we think. It can become a tool to preserve a critical distance to mediation in order to have room to interpret the experience.

This thesis offers an expressive vocabulary and techniques for virtual reality, which, ideally, will yield to new forms of expressivity and ways of seeing. In the case of virtual reality, a medium touted for its mimetic powers, the lure of realism is stronger than ever. However, this constrains the creative and non-mimetic uses of the medium. My project in this thesis is closer to the project of the modernists; my position is akin to those of Rudolf Arnheim and Sergei Eisenstein, early theorists of cinema who break from the assumptions of film's realism in order to create an expressive vocabulary peculiar to the medium itself. In the case of virtual reality, with added technological elements such as responsiveness, embodied hardware and spatial capture & navigation, verisimilitude takes on new meanings and new stakes. It is constructed through a more complicated network of relations; such as the indexical data of the physical world being governed by a more or less arbitrary rule system. However, because so many of these changes are tied to technological innovation, a modernist inspired analysis that directly accounts for the medium itself is productive.

There is a common trajectory in the works discussed throughout the thesis. *In the Eyes of the Animal, Timescrubbing, Phantom,* along with other works, all try to break away from the realism, which is presented and marketed as the essential trope of the medium. Even though virtual reality offers a heightened sense of presence and immediacy, these are not the only aspects that are

significant about the medium. As I have shown, presence and immediacy factor into work created in many mediums beyond virtual reality. This thesis has attempted to discover and enumerate the qualities of virtual reality as a medium that are unique and deserve our attention. The construction of virtual reality works for expressive non-mimetic uses allows for a more promising future for the medium. Research in the field of embodied cognition, as well as advancements in 3D capture, virtual reality hardware and experiences suggest that the contingent, responsive, and embodied aspects of virtual reality will continue to be central to our future, and are likely to change and impact our relationship to knowledge and the world around us.

Bibliography:

Arnheim, Rudolf. *Film as Art*. Berkeley: U of California, 1969.

Au, Wagner James. "VR Will Make Life Better-Or Just Be an Opiate for the Masses." *Wired.com*. Conde Nast Digital, 25 Feb. 2016. Web. 04 May 2016.

Aylett, R. *Emergent Narrative, Social Immersion and "Storification"*. Proceedings Narrative and Learning Environments Conference NILE00, Edinburgs, Scotland. 2000.

Bazin, André. *What is cinema?*. Berkeley: University of California Press, 2005.

Benjamin, Walter. *The Work of Art in the Age of Its Technological Reproducibility, and Other Writings on Media*. Cambridge, MA: Belknap of Harvard UP, 2008.

Bogost, Ian, Simon Ferrari, and Bobby Schweizer. *Newsgames: Journalism at Play*. Cambridge, MA: MIT, 2010.

Bolter, J. David, and Richard A. Grusin. *Remediation: Understanding New Media*. Cambridge, MA: MIT, 2000.

Bordwell, David, Janet Staiger, and Kristin Thompson. *The Classical Hollywood Cinema: Film Style & Mode of Production to 1960*. New York: Columbia UP, 1985.

Boubekeur, Tamy. "Mesh Reconstruction from a Point Cloud". *Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality*. Eds. Magnor, Marcus A., Oliver Grau, Olga Sorkine-Hornung, and Christian Theobalt. Boca Raton, FL: CRC Press, 2015.

Boyacioglu, Beyza. "Embodiment and Mimetic Representation in Non-Fiction Virtual Reality Experiences". 2015. TS.

Burdette, Matt. "The Swayze Effect." *Story Studio Blog*. N.p., 18 Nov. 2015. Web. 04 May 2016.

Canales, Jimena. *A Tenth of a Second: A History*. Chicago: U of Chicago, 2009.

Casares, Adolfo Bioy. *The Invention of Morel: And Other Stories (from La Trama Celeste)*. Austin: U of Texas, 1964.

Colomina, Beatriz. "X-Screens: Röntgen Architecture." *E-flux*. N.p., 2015. Web. 04 May 2016.

Crary, Jonathan. *Techniques of the Observer: On Vision and Modernity in the Nineteenth Century* Cambridge, Mass.: MIT Press, 1990.

Dancyger, Ken. *The Technique of Film and Video Editing: History, Theory, and Practice*. New York: Focal, 2011.

Deleuze, Gilles. *Cinema*. Minneapolis: University of Minnesota, 1986.

Dorsky, Nathaniel. *Devotional Cinema*. Tuumba Press, 2005.

Eisenstein, S. and J. Leyda. *Film form: Essays in Film Theory*. New York: Harcourt, 1949.

Elsaesser, Thomas. "Pushing the contradictions of the digital:'virtual reality'and 'interactive narrative'as oxymorons between narrative and gaming." *New Review of Film and Television Studies* 12.3*, 2014, 295-311.

"FARO Focus3D." *Laser Scanner Faro Focus*. N.p., n.d. Web. 14 Dec. 2015.

Fish, Stanley E.. "Interpreting the "variorum"". *Critical Inquiry 2.3* (1976): 465–485. Web.

Foucault, Michel. *The Order of Things: An Archaeology of the Human Sciences*. New York: Vintage Books, 1973.

Galloway, Alexander R. *The Interface Effect.* Cambridge, UK; Malden, MA: Polity Press, 2012.

———. *Gaming: Essays on Algorithmic Culture*. Minneapolis: University of Minnesota Press, 2006.

Gibson, James J. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979.

Grau, Oliver. *Virtual Art: From Illusion to Immersion*. Cambridge, Mass.: MIT Press, 2003.

Gumbrecht, Hans Ulrich. *Production of Presence: What Meaning Cannot Convey.* Stanford, Calif.: Stanford University Press, 2004.

Hansen, M. B. N.. "Algorithmic Sensibility: Reflections on the Post-Perceptual Image". *Post-Cinema: Theorizing 21st-Century Film*. Eds. S. Denson and J. Leyda. REFRAME, 2016. Web.

Harrell, D. Fox. *Phantasmal Media: An Approach to Imagination, Computation, and Expression.* Cambridge, Massachusetts: The MIT Press, 2013.

Herzog, W.. *Of Walking In Ice: Munich-Paris, 11/23 to 12/14, 1974*. New York, Tanam Press. 1980.

Hoelzl, I. and R. Marie. *Softimage: Towards a New Theory of the Digital Image*. Intellect (UK), 2015.

Hoelzl, Ingrid. "The Operative Image - an Approximation." *MediaCommons*. N.p., 3 Feb. 2014. Web. 04 May 2016.

House, Patrick. "Werner Herzog Talks Virtual Reality." *The New Yorker*. 12 Jan. 2016. Web. 25 July 2016.

Johnson, Eric. "Virtual Reality Is 'The Last Medium,' Says Filmmaker and Vrse CEO Chris Milk (Q&A)." *Recode*. N.p., 01 Oct. 2015. Web. 04 May 2016.

Johnson, Mark. *The body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press, 1987.

———. *The meaning of the Body: Aesthetics of Human Understanding*. Chicago: University of Chicago Press, 2007.

Jenkins, H.. "Game Design as Narrative Architecture". *First Person: New Media as Story, Performance, and Game.* Eds. N. Wardrip-Fruin and P. Harrigan. Cambridge, Mass.: MIT Press, 2004.

Kolb, Andreas and Pece, Fabrizio. "Range Imaging". *Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality.* Eds. Magnor, Marcus A., Oliver Grau, Olga Sorkine-Hornung, and Christian Theobalt. Boca Raton, FL: CRC Press, 2015.

Lakoff, George and Mark Johnson. *Metaphors We Live By*. Chicago: University of Chicago Press, 1980.

Manovich, Lev. *The Language of New Media*. The MIT Press, 2001.

Marsh, Calum. "The Shot of the Year." *The Dissolve.* N.p., n.d. Web. 04 May 2016.

Milk, Chris. "How Virtual Reality Can Create the Ultimate Empathy Machine." *TED.* N.p., n.d. Web. 04 May 2016.

Nichols, Bill. *Introduction to Documentary*. Bloomington, IN: Indiana University Press, 2001.

Packer, Randall, and Ken Jordan, eds. *Multimedia: From Wagner to Virtual Reality*. New York: Norton, 2002.

Prince, Stephen, and Wayne E Hensley. "The Kuleshov Effect: Recreating the Classic Experiment" *Cinema Journal* 31.2, 1992: 59-75

Rees, A. L. *Expanded Cinema: Art, Performance, Film*. London: Tate Gallery Pub., 2011.

Ropero-Miller, Jeri. *Landscape Study on 3D Crime Scene Scanning Devices*. Rep. N.p.: National Institute of Justice, n.d. Print.

Ruiz, Raúl. *Poetics of Cinema*. Paris: Éditions Dis Voir, 1995.

Schroeter, Jens. *3D: History, Theory, and Aesthetics of the Transplane Image*. New York: Bloomsbury Academic, 2014.

Shaviro, Steven. *The Universe of Things: On Speculative Realism*. Minneapolis, MN: University of Minnesota Press, 2014.

Smith, Merritt Roe, and Leo Marx, eds. *Does Technology Drive History?: The Dilemma of Technological Determinism*. Cambridge, MA: MIT, 1994.

Sobchack, Vivian Carol. *The Address of the Eye: A Phenomenology of Film Experience.* Princeton, N.J.: Princeton University Press, 1992.

———."The Scene of the Screen: Envisioning Photographic, Cinematic, and Electronic 'Presence'". *Post-Cinema: Theorizing 21st-Century Film*. Eds. S. Denson and J. Leyda. REFRAME, 2016. Web.

Steyerl, Hito. "Ripping Reality: Blind Spots and Wrecked Data in 3D." *Eipcp*. N.p., n.d. Web. 04 May 2016.

Sutherland, E. Ainsley. *Staged Empathy: Empathy and Visual Perception in Virtual Reality Systems*. Master's Thesis, MIT, 2015.

Szeliski, Richard. "Foreword". *Digital Representations of the Real World: How to Capture, Model, and Render Visual Reality*. Eds. Magnor, Marcus A., Oliver Grau, Olga Sorkine-Hornung, and Christian Theobalt. Boca Raton, FL: CRC Press, 2015.

Taylor, Richard. *Film Propaganda: Soviet Russia and Nazi Germany*. London: I.B. Tauris, 1997

Thiel, Tamiko. "Where Stones Can Speak: Dramatic Encounters in Interactive 3D Virtual Reality". *Third Person: Authoring and Exploring Vast Narratives*. Eds. Harrigan, P. and N. Wardrip-Fruin. Cambridge, Mass.: MIT Press, 2009.

Tortum, Deniz. "How Virtual Reality Technology Is Changing Documentary Filmmaking." *Indiewire*. N.p., 31 Jan. 2015. Web. 04 May 2016.

Uricchio, William. "The Algorithmic Turn: Photosynth, Augmented Reality and the State of the Image". *Visual Studies* 26:1 (2011): 25-35

Varela, Francisco J., Evan Thompson, and Eleanor Rosch. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT, 1991.

Virilio, Paul. *The vision machine*. Bloomington: Indiana University Press, 1994.

Wardrip-Fruin, Noah, and Nick Montfort, eds. *The New Media Reader*. Cambridge, MA: MIT, 2003.

"WHAT IS DIGITAL SYNESTHESIA? | Digital Synesthesia." *Digital Synesthesia*. N.p., n.d. Web. 04 May 2016.