November 2021

Experiencing Cinematic VR: Where Theory and Practice Converge in the Tribeca Film Festival Cinema360

John V. Pavlik
Rutgers, the State University of New Jersey, jpavlik@comminfo.rutgers.edu

Experiencing Cinematic VR: Where Theory and Practice Converge in the Tribeca Film Festival Cinema360

John V. Pavlik, *Rutgers*, *The State University of New Jersey*

Conference Paper (Faculty)

Abstract

Cinematic virtual reality (VR) production has reached enough capacity to support a festival. This paper offers a theoretical framework of VR narrative structure to critically examine one such festival in cinematic VR. The spotlight here is on the fifteen entries in the 2020 Tribeca Film Festival Cinema360. Findings suggest that although the field of cinematic VR has advanced substantially in recent years in terms of narrative design and user experience, there is still a considerable distance for VR storytellers to travel to fully utilize the nature and potential of the developing medium of virtual reality.

Keywords: cinematic virtual reality, immersion, media, storytelling, 360 video

The COVID-19 Context

The widespread stay-at-home orders imposed in response to the COVID-19 pandemic have given people an increased incentive to find novel and engaging activities in which to participate at home. Among these unusual at-home experiences is virtual reality (VR) or other immersive media. Immersive media content experiences can take a variety of forms, including video games, art, journalism and immersive cinema. This paper examines the domain of immersive cinema, or cinematic VR as it is more widely known. Based on a theoretical framework that outlines the nature of narrative design in immersive media, we critically examine the fifteen entries in the 2020 Tribeca Film Festival Cinema360 competition.

Literature Review

In 1901, *Wizard of Oz* author L. Frank Baum imagined electronic glasses that enabled the wearer to see a virtual "character marker" displayed on a fellow human being. Since then, musings about augmented reality (AR) and virtual reality (VR) have moved from the realm of science fiction into the domain of commercial fact. Computer scientist, mathematician and artist Jaron Lanier coined the term virtual reality (VR) in 1987, the same year he introduced the first commercial VR headset (Virtual Reality Society, 2017). In 1990, Boeing engineer Tom Caudell coined the term augmented reality (AR) as a new approach to plane design and production (Vaughan, 2019).

Höllerer, Feiner, and Pavlik (1999) soon developed an AR-based form of storytelling called the Situated Documentary. This form presented non-fiction, news content of past events (e.g., photos, video) displayed on a see-through heads-worn display, with digital content overlaid onto the actual physical location where events had occurred. This ARsystem used geo-location enabled by a combination of Global Positioning Satellite (GPS) and WiFI. Audio playback via headphones provided a narration track and other acoustical elements including audio recordings from past events being re-presented. Omnidirectional images (360 degree) were displayed on the headworn display and the user could turn their head in any direction to look about, or could physically move about in a threedimensional space to move to other locations and access additional AR content layers.

Still, the public embrace of these emergent digital media forms has been ambivalent at best. The initial marketplace enthusiasm that in 2013 first welcomed Google's AR headset dubbed "Glass" quickly turned into marketplace resistance and even revulsion due to privacy concerns. Some even disparaged the wearers as "glassholes" (Honan, 2013). Some three years later, Facebook's Oculus Rift and HTC's Vive VR HMDs entered the global marketplace with some initial fanfare, as many thought VR had finally come of age (Gownder, Voce, Mai, and Lynch, 2016; Lang, 2019). But that promise waned amidst the high initial cost and technical complexity of the platforms, and especially the growing public realization that there was little quality content available, and most of that was limited to the domain of games.

The past three decades have seen relatively slow growth in the public market for AR and VR. "AR and VR have been afflicted by the small size of their respective markets for a long time now. Last year (2019), the global shipments of VR headsets totaled 5.7 million, which compares pretty poorly to the over 1.3 billion smartphones shipped in the same amount of time" (Leprince-Ringuet, 2020).

Immersive media in the form of augmented reality (AR) and virtual reality (VR) may be poised for significant growth in public adoption and usage, however. Smartphones are an increasingly effective platform for AR, and there are more than 3.8 billion of those in operation worldwide (Statista,

2020). A wide variety of organizations have been creating AR experiences designed for smartphones. Among these are news organizations such as *The New York Times*, which since the 2018 winter Olympics in Seoul, South Korea, has been creating a growing number of news stories utilizing AR for the iPhone, with others creating AR experiences for Android OS smartphones.

Likewise, there are increasing examples of media and other enterprises creating immersive content experiences designed for wearable platforms, such as the Oculus and the HTC Vive head-mounted displays (HMDs). From the BBC to USAToday, a growing number of news media organizations around the world are producing immersive news content. The Sundance Film Festival gave its top award to the immersive journalism production "Hunger in Los Angeles" (Gilmour, 2012; De la Peña, 2017; De la Peña, et al., 2010).

Complementary to the immersive news arena, at least a half dozen production studios have produced cinematic VR. Cinematic VR provides sight and sound narratives analogous to traditional cinema, but designed as enveloping experiences meant for HMDs (i.e., the motion pictures and sound are omnidirectional, taking in the wearer's entire field of view and acoustic experience). These cinematic VR studios include Occupied VR (Toronto, Canada), Facebook's Oculus for Good productions, Thirteen Floor Experiential Media Group (Seoul, South Korea) and Funique VR Studio (Taipei City, Taiwan).

Moreover, "Recent research released by IDC indicates that despite problems in the enterprise created by the COVID-19 outbreak this year, it is likely to be a good year for AR and VR vendors (Roe, 2020). The "Worldwide Quarterly Augmented and Virtual Reality Headset Tracker" (2020) shows that although the market for AR and VR may decline by 10.5% in the first quarter of 2020 followed by a further decline of 24.1% "as a result of supply chain disruptions created by the current global situation, there will be a rebound later in the year with growth of 23% globally in the market, up by nearly 24% on last year" (Roe, 2020).

Moreover, other data confirm that users of these immersive platforms are increasingly connecting with the online sources of immersive content, including the popular game platform Steam. "The latest figures from January (2020) show the single biggest leap yet in monthly-connected VR headsets on Steam, now with 1.31% (+0.22%) of the population having headsets connected to their computers. With a year-over-year growth of 43%, this beats out the previous record set in January 2019 when monthly-connected headsets jumped by 0.11%, reaching 0.91% for the first time" (Lang, 2020). In April 2020 about 2% Steam users connected to the platform via a VR headsets, exceeding the previous high of 1.5% by one-third (Collins, 2020).

In an investigation of the interactive documentary, a web-based precursor to immersive cinema, Pavlik and Pavlik (2015) outlined a theoretical framework that identified "key dimensions of storytelling in the interactive documentary, a media form using digital, networked and mobile media." They state that among the relevant dimensions are the concepts of interactivity, multiple media, and dynamic narrative or non-linear story structure in which an active user controls or even shapes the story pathway with no common single beginning, middle or end; as such, users do not necessarily all experience the same story (Murray, 2017). Traditional analog media tend to feature linear narrative story structure, which gives the audience a common experience (Pinault, 1992; Mittell, 2015). Audiences in traditional narrative forms tend to engage stories in a relatively passive form (McQuail, Blumler, and Brown, 1972).

Pavlik (2018) has refined the digital narrative framework to apply to the platforms of AR and VR, what he calls part of "experiential media." Experiential media allow users to experience virtual phenomena as if present or participants in the content, such as cinematic VR. In this experiential media framework, he identifies six narrative dimensions, including: 1) interactivity, 2) immersion, 3) multi-sensory communication, 4) data-driven artificial intelligence (AI), 5) first-person perspective and 6) natural user interface (NUI). We apply this six-dimensional conceptual framework in the current study to the study of cinematic VR.

Interactivity is defined here as a form of exchange or communication dialog between users and the content experience or with others simultaneously (or asynchronously) engaged in the experience, whether remotely or in physical proximity. In experiential media platforms, users have the ability engage in communication or information exchange with others. They also have the ability to actively control, navigate or otherwise participate in the virtual experience, alter the path of the narrative or even input their own communications. A growing amount of research is examining the nature or user engagement in AR and VR and the role of interactivity in the content designed for AR, VR and related platforms (Hassenzahl, 2013).

Immersion refers to the envelopment of the user, whether visually, aurally or via other senses (e.g., haptic). Critical to understanding the nature of the immersion is what developers call the degrees of freedom within the immersive environment. Omnidirectional, or 360-degree video, has three degrees of freedom (DOF). That is, the user can look in any direction, up/down, left/right, forward/back. This can give the user a sense of being enveloped in a virtual space or a 360-degree image or video of reality or animation. However, the user cannot move about in a three-dimensional (3D) space. That is, the user occupies a fixed space, but can look about. They move about in the virtual environment only if the content developer has encoded movement, such as by moving the camera or programming

movement. Such immersive content offers a step toward fully immersive VR. But for a more fully immersive VR experience, the content developer must encode six DOF. With 6 DOF, the user not only can look about but they can move about, traveling forward/backward, up/down, left/right. Such ability to move can give the user much more complete sense of immersion in VR. Creating immersive content with six DOF requires substantially more programming and production than simply video or audio recording 360-degree fields of view or sound. Conventional video occupies 90-135 degrees of the user's field of view, leaving most of a person's peripheral vision outside the visual media frame. Therefore, the image or video does not fully occupy the individual's visual frame of perception, and possibly reduce psychological immersion. Likewise, 3D audio that provides spatial sound (with 3D integrity in the virtual environment), not just stereophonic, requires a more complex array of microphones that can also maintain geolocative synchronicity.

Multi-sensory communication typically takes the form of visual and aural. But with newer VR platforms haptic, or tactile, user engagement is also available. Taste and smell are also possible in a virtual environment although most current VR platforms do not include technology to enable user taste or smell experiences.

Data-driven AI takes the form of advanced algorithms (programmed instructions or coding) and sensors that track user actions (e.g., gestures, eye movement) and support the generation of an experience with near-zero latency. This means the delay between a user action and a response from the system or virtual experience is imperceptible to the user. It makes the virtual experience seem more real; the user's action, such as tapping on a digital object and getting a reaction, seems to be causally related to the responses in the system and gives the user a sense of agency in the virtual world.

First-person perspective means the user enters the virtual experience as if present as a participant or virtual witness to events or experiences. AR and VR fall along a continuum of virtuality as theorized by Milgram, Takemura, Utsumi and Kishino (1994). AR blends users' real-world and virtual experiences, while VR offers synthetic experiences or computerized experiences that replace users' direct contact with the physical world. Immersive, participatory formats are especially amenable to a 1st person perspective because they enable the user to do things, to take actions and see or otherwise experience the consequences.

The natural user interface (NUI) means that the user engages the system and interacts using intuitive means of communication, including voice, gesture, touch or gaze (Marcus, 2015). This enhances the user's sense of presence within the virtual environment and enables participation without the need for training, literacy or other more technical means of interaction and experience navigation (Zahoric and Jenison, 1998).

Combined, these dimensions enable the user to experience virtual phenomena as if present in virtual environments, thus the label "experiential media." Bailenson, director of Stanford's Virtual Reality (VR) laboratory, describes VR as "experience on demand" (2018:1). A haptic interface could enable the user to feel virtual objects (Parisi, Paterson, & Archer, 2015). Research indicates that the combined dimensions outlined above can facilitate telepresence (Sundar, Kang, & Oprean, 2017; Archer & Finger, 2018).

Telepresence, often shortened to "presence" in an immersive media context, refers to "touch and feel at a distance" (Slack & Wise, 2005, p. 38) and occurs when a medium of communication disappears to the user (Kim & Biocca, 1997). This apparent absence or disappearance of the medium constitutes presence in immersive media such as AR and VR and allows the user to seamlessly enter into a virtual world or experience and engage more immersively in a psychological sense as well, as McLuhan suggested when he posited that the medium is the message (1964).

The Current Research Question

By 2020, the quality of experiential media platforms improved dramatically and the cost had fallen to a just few hundred dollars from several thousand. The installed user base for AR and VR reached a potential critical mass. Moreover, the volume of available immersive content experiences began to offer users a potentially compelling reason to enter into the so-called metaverse, a concept introduced by Stephenson in his science fiction classic *Snow Crash* (1992). The metaverse, referring to a convergent networked digital environment blending virtually enhanced physical reality with physically persistent virtual spaces, seems to have advanced substantially with the stay-at-home orders implemented during the COVID-19 crisis (Smart, 2007). Record numbers of persons around the world immersed themselves in online social platforms such as Zoom, Twitter and Facebook, and engaged increasing amounts of time in immersive media forms ranging from immersive games to VR concerts. One notable example is a virtual concert hosted on the extraordinarily popular global video game platform called Fortnite from Epic Games, based in North Carolina. As of 2019, Fortnite had 250 million players worldwide, and the game generated an estimated $1 billion in revenues monthly, mostly from in-game purchases by the players. On April 24, 2020, Epic hosted a virtual concert by hip-hop artist Travis Scott, with a record 12.3 million concurrent viewers (Rogers, 2020). The previous record for a virtual concert was DJ Marshmello's performance February 2019 with 10 million concurrent viewers. Scott's *Astronomical* captured 27.7 million total views across the five events that ran through 27 April. In the concert, a virtual Scott was scaled to more than 100 feet in height, and teleported virtually across the universe as he performed his popular music, featuring themed in-game items for purchase and a psychedelic road trip through space.

In this paper we journey into one part of the metaverse, that of cinematic VR, the domain of VR storytelling and experiences. We ask how and to what extent are content creators of immersive cinema utilizing the unique qualities of the experiential media platform of VR in the 15 entries in the TriBeCa VR film festival titled Cinema360 and held in 2020.

Other film festivals are making their cinematic VR entries available online on the Oculus Quest platform, as well, among them SXSW which featured "After the Fallout" (20 May 2020).

Method

Because of the COVID-19 pandemic, the 2020 Tribeca festival was moved entirely online, where public screening of the Cinema 360 entries was available on-demand and at no cost on Oculus TV 17-26 April 2020 (Stein, 2020). To view, or more precisely experience, the entries in an immersive format required the user to don an Oculus headmounted display (HMD), such as the Oculus Quest, which the author used. The HMD also required a WiFi connection to access and download the content. It should be noted that the festival entries examined here are not presented as full VR. They are entered in a festival category labeled "360", indicating that the immersiveness is framed in the context of three DOF, not six. There are other categories designated VR, with presumably six DOF content.

The analytical framework used here is based on identifying whether and the extent to which each entry employs the six dimensions of experiential media storytelling. This means 1) whether a video employs interactivity and in what form, 2) the user of immersion, including three (or potentially six DOF) visually, three-dimensionality imagery, and audio envelopment, 3) the use of multisensory engagement (sight, sound, haptics), 4) the role of AI and/or data to provide a zero-latency adaptable user experience, 5) the use of first or third person perspective and which is dominant temporally, and 6) the use of natural user interface (voice, touch, gaze or other).

The fifteen entries in the 2020 Tribeca Film Festival Cinema360 competition serve as the data source for this paper and are listed in Figure 1. The author experienced each entry (i.e., viewed immersively) twice during their one-week of their availability on OculusTV 17-26 April 2020. Each entry was evaluated in terms of its utilization of the six narrative dimensions of experiential media outlined above and summarized in figure 1. All fifteen productions were recorded and posted on the OculusTV, where the author downloaded them onto his Oculus HMD. He viewed each alone, with no social interaction or social media. Non-immersive conventional video trailers of some of the productions are on Vimeo (2020). Although some of the author's study measurements are quantitative, this study is essentially an impressionistic one. Results cannot be generalized from the reactions and

coding reported here. But it is hoped the findings will be suggestive of the value and relevance of the key concepts of experiential media as they are manifested in this current generation of cinematic VR.

To best understand the nature of the immersive viewing experience, it is worth describing in some detail the Oculus Quest of 2020; see images 1 and 2. An immersive, head-worn device, the Quest is markedly better than the Google cardboard of 2015, which required the user to fold a cardboard cutout into a headworn display into which the user inserted their smartphone. The viewing experience was similar to a Sawyer's View-Master display. The Quest of 2020 is made of molded plastic, with built-in computer and WiFi technology, spatial speaker playback. The device is adjustable to the user's head size and eye-sight focal length. The display is of substantially higher resolution, comfort, and omnidirectional field of view with immersive audio quality and haptics. In a relative to television sense, the immersive platforms of 2015 were akin to an early B&W 12" TV sets with mono-phonic sound of 1940s which were often viewed head-on through a magnifying glass, in contrast to 2020's 60" ultra high definition color TV displays with surround sound. There are similarly dramatic differences in content, too.

"Launching back in 2019, Oculus Quest is the all-in-one VR gaming system," wrote one industry observer (Vane, 2020). "There are no cables, and no PC is required at all (previous HMDs required a cable connected to a PC, greatly limiting the immersive experience). It's an entirely wireless experience – except when charging the headset, of course. Two Touch controllers are even included with the headset – one for each hand. All you need to do is buy some games or apps." The system also allows the user to employ their own hands as the controllers.

Further, the "Oculus Quest handles motion tracking with four built-in ultra-wide-angled sensors that track your exact position in real-time. Unlike the entry-level Oculus Go, the Quest provides 6 DOF motion tracking. This means your movement will be recognised around a room versus just looking and pointing." As to the quality of the visual display, the Oculus features an OLED display with "resolution of 1600×1440 per eye while incorporating a lens spacing adjustment to help maximise visual comfort. Oculus has also improved the built-in audio, so you get higher-quality, immersive sound with even deeper bass."

Using the Oculus also requires certain safety precautions to protect the user, who once inside a virtual environment can easily forget where they are in the physical world and could be injured by physical objects not visible in the virtual space. As a result, the Oculus has a built-in protective virtual field (with a minimum 5'x5' space), called "the Guardian" that the user calibrates before using any app or watching any cinematic VR. The "Guardian technology carries over from the Oculus Rift with built-in outwards-facing

monochrome cameras used to set up the boundaries of your play area to avoid injury. The system can recognise and store multiple play areas, so you can quickly move VR from room-to-room or house-to-house" (Vane, 2020). The Oculus also recommend to the user the taking of breaks from VR every 30 minutes.

As to the haptics or touch interface, "Two (handheld) Touch controllers are bundled with the headset, so there's no need to purchase expensive accessories. The controllers bring your hands and gestures into VR so you can easily interact with virtual worlds. Your slashes, throws and grabs appear in VR with intuitive, realistic precision. The controllers feature two buttons, a thumbstick and trigger for merging traditional controls with your physical gestures" (Vane, 2020).

Findings

Cinema360 grouped the fifteen international entries (only three are U.S.A. productions) into four categories, *Dreams to Remember, Seventeen Plus, Kinfolk,* and *Pure Imagination,* with three or four entries in each. This analysis focuses on each production's use of narrative techniques, particularly the qualities of experiential media. Figure 1 summarizes the findings. We begin our discussion with the four entries in *Dreams to Remember*.

[INSERT FIGURE 1 ABOUT HERE.]

1st Step is a non-fiction immersive documentary produced in Germany. It tells the story of the Apollo 8 lunar mission of 1969 as well as more recent space travel by NASA astronauts. Some of the narrative is in the 3rd person when the story focus is the astronauts themselves. The user, however, is largely engaged in the 1st person as an active participant rather than as a passive viewer or audience member. The experience begins by transporting the user into an immersive space through a portal like 3D doorway. The user soon ascends via an elevator rising to the Apollo 8 space capsule. The experience uses actual NASA film footage and images presented in 360-degree format. The user can look in any direction at any time. Much of video features 3D visual imagery with compelling depth of field. Spatial audio plus English-language narration accompanies the user on their journey to the surface of the moon. Although the historical narrative is linear, non-interactive, non-haptic and features only 3 DOF, the high-resolution, color and omnidirectional visualization and spatial sound generates a strong sense of presence for the user and brings a high level of verisimilitude. After experiencing the narrative these were the author's immediate thoughts: I've just journeyed to the moon and back and it was inspiring.

Rain Fruits is a South Korean production blending animation and live action into a 3D 360-degree 3 DOF linear fictional narrative. There is no interactivity or haptic engagement, but original music and narrated, English-language audio surrounds the user as they travel through a 3rd-person story of a Myanmar migrant as an expatriate worker in South Korea. The social justice themed narrative offers a critique of capitalism, although the migrant's journey transforms from one of initial frustration and homesickness into eventual happiness, at least in a professional context as he gains entry into the world of professional engineering. One notable scene briefly transitions from 3rd person to 1st person viewpoint, and the user quickly enters as though running into the subway entrance. A side effect of this is the user may feel some motion sickness (or VR sickness), if they are not well acclimated to the VR platform. This experience highlights the need for the user to build up tolerance for wearing an HMD.

Dear Lizzy is a colorful animated 3D and 360-degree environment in which the user watches as a woman shares her English-language narrated and musical journey of missing Lizzy, her dear friend. The user virtually travels through various fanciful land, sea and space scapes. In a few short sequences, the user sees the through world via a 1st person view, but most of the experience is told in the 3rd person voice. The journey is linear and the user can look around with 3 DOF in this brief American production.

Forgotten Kiss is a production from Finland blending live action and animation in a complex and colorful 3D, 360-degree fantasy. It is an adaptation of a Russian folk tale. Narrated in English, the spatial audio features a techno music background in dreamlike fantasy. In a combination of 1st and 3rd person perspective, the user travels through the linear narrative, although they are drawn to look in various directions through audio cues, such as a voice from behind, or an unusual musical note left or right.

Category **Seventeen Plus** features four productions. We begin our analysis with a narrative that focuses on suicide. Reminiscent of the classic CBS TV show *The Twilight Zone* or the contemporary Netflix series *Black Mirror*, *A Safe Guide to Dying* is a Greek production that offers a largely 3rd-person linear narrative of Linus Fielding, a middle aged man who gradually comes to realize he exists in a virtual world. Somewhat evocative of HBO's *Westworld*, Linus discovers he is trapped in this virtual world. Becoming increasingly alarmed and depressed, he tries to commit suicide, but finds that cannot kill himself. The compelling narrative uses 360-degree visualizations effectively to advance narrative flow, incorporating some 3D effects. Spatial audio directs the user's occasional 1st person perspective to look in various directions. The live action narrative is occasionally nonlinear in its flow, but only under the control of the production and not the user.

Black Bag, a Chinese production, is a dark, animated, hand-painted, and abstract 360-degree narrative of an urban bank robbery, suggestive of the

popular video game, *Grand Theft Auto*. The depiction of only white males in the narrative limits its accessibility for diverse users. Some occasional 3D imagery enhances the 3rd-person narrative.

The Pantheon of Queer Mythology is a Spanish production that engages the user in immersive, 360-degree imagery, with 3 DOF and compelling 3D visual fantasy depictions of Queer culture. The linear animated narrative makes effective use of scale. Placing the user in the role of a visitor to an imaginary museum, the user travels much of the story in the 1st person. The narrative transitions cleverly by transporting the user from the ordinary to the mythological Queer world beyond. Some sexually explicit imagery populates the story, which explores the diversity of human identity. With a broad range of color, sound and fantastical oversized imagery, this production offers indirect interactivity via a companion website.

Inspired by Goya's *Saturn*, *Saturnism* is a French production. This linear 360-degree animation begins in near darkness as the user hears a strange crunching sound. The perspective is 1st-person, and the user begins to realize they are in a cave. As the scene before them becomes more visible, the user realizes they are witnessing the giant god Saturn devouring a human man. The experience seems uncomfortably real when Saturn's large pointed teeth seem to clamp down on the user's head. The author found himself wanting to yank off the headset. The user is about the size of Saturn's hand and feels as vulnerable as Gulliver in Brobdingnag, the land of giants in Swift's classic tale. Spatial audio effects range from the sounds of footsteps in the cave to Saturn's eating. The sense of presence is powerful and terrifying; the author had to keep reminding himself it wasn't real.

Category ***Kinfolk*** includes four productions. We start with an analysis of an Ethiopian production, *Ferenj: a graphic memoir in VR*. *Ferenj* means foreigner. The 3D 360-degree animated narrative explores mixed race identity through the lens of a first-generation Ethiopian-American. The story is presented largely in a 3rd-person perspective, and the user hears in English narration the protagonist tell the story of her conflicted identity. "I grew to hate the fact there is white in me," she states. The 3 DOF linear narrative depicts cultural scenes visually and through a combination of Ethiopian and American music. One compelling part of the story is the narrated memories of the woman conversing with her parents and friends. A few moments switch the user into a 1st-person viewpoint during this dialog.

The Inhabited House is an auto-biographical production by Argentine director Diego Kompel. Kompel narrates the story and shares many of his own still photos and home movies interspersed to create the immersive 360-degree non-fiction personal documentary set inside home of his grandparents. There are no 3D components and the story is largely linear in a temporal sense, and the user has no control of the narrative flow or movement of the story. Yet, the experience has warmth as a combination of spoken word English and

Spanish blend together with ambient sounds, music and singing from the past.

On its face, *Home* is perhaps the simplest of the entries in the Cinema360 competition, and also the most complex. Produced Taiwan, R.O.C., the production is a virtual visit inside a Chinese home in Taiwan (R.O.C.). Employing an entirely 1st-person point of view, the user immediately assumes the role of the grandmother, whose multigenerational family has come for a visit on a Chinese holiday. With quality spatial audio, the conversation is in Chinese with subtitles. Within the narrative, a family member walks up to you, the user, grandmother, and leans in, looking you directly in the eye. The feeling is uncanny and seems real. When he speaks to you, the feeling of a need to respond is profound. The live action, entirely shot within an actual two-room home in Taiwan is highly authentic down to fine details, including items of food, clothing and a Chinese program playing on the TV. As family members begin preparing to depart, they first come close to you, look into your eyes, address you and express their love and regrets that they must leave, but they will be back. Suddenly, you begin to move and, if you glance down, you see you are in a wheelchair, and behind you is your nurse pushing you toward the wide, open patio entrance to your home, where family members have gathered for a family photograph. Produced by Funique VR studio, with director Kidding Hsu, the live action production features a large cast of more than a dozen young and old actors. Throughout, the user can't help reacting by talking, responding, nodding, waving. At times forget it's easy to forget the actors on screen can't hear you; the simulated interaction is deeply felt. And it's hard to not feel sad when they leave you alone. Shot in 360-degree video, there are a few well-chosen 3D movements such as that of the rolling wheel chair. *Home* seems simple but delivers high verisimilitude. It's the difference between watching people and being in the room with them. Utilizing the VR Film Lab, the production features 8K (i.e., ultra high definition) video (with stereo audio).

Home evokes a concept from aesthetics known as "the uncanny valley," an "hypothesized relationship between the degree of an object's resemblance to a human being and the emotional response to such an object." Humanoid objects that "imperfectly resemble actual human beings provoke uncanny or strangely familiar feelings of eeriness and revulsion in observers" (Mather & Reichling, 2016). The term "valley" refers to "a dip in the human observer's affinity for the replica, a relation that otherwise increases with the replica's human likeness." In the case of VR, this uncanniness is acute when actors in the scene approach the user, look them in the eye, and apparently interact with them. The apparent reality of the experience makes the user almost forget it is only virtual.

This uncanniness is amplified by the breaking of the 4th wall notion endemic to theater and film. In these performance arenas, the actors exist in a world

bounded by three walls, and separated from the audience by a so-called 4th wall. Since the time of Shakespeare, productions occasionally have utilized the narrative device of breaking the illusory 4th wall to engage the audience directly by having an actor look at and speak to the audience rather than another actor. In cinema, this technique sometimes has been used to enable an actor to speak directly to the audience. In TV shows, hosts often use this technique when introducing a show, look into the camera lens, and speak directly to the audience (e.g., director Alfred Hitchcock would do this in the opening and closing of his show, *Hitchcock Presents*, in the 1950s). Reporters in the field and news anchors in the studio will typically look directly into the camera lens and speak to the audience, temporarily stepping out of the 3rd-person voice and into a 2nd-person perspective. Research suggests the technique of direct eye contact can exert multiple psychological impacts, including increasing self-consciousness and audience trust in the news presenter (Jarrett, 2016). In *Home*, this technique creates a simulated interaction between the user and the rest of the characters, who are analogous to humanoid robots within the VR production, making the experience seem all the more real to the user.

The category of **Pure Imagination** features the final four productions. Supported by the Oculus *VR for Good program* (part of Facebook, which owns the Oculus), *Lutaw* is an animated production from the Philippines. The 360-degree narrative begins with the user submerged underwater in the Philippines sea. 3D visual effects are compelling, with the user briefly in a 1st person perspective, experiencing a close encounter with a shark and other sea life. After swimming to shore and the story transitions to the 3rd person, the user observes the travails of a young girl and her even younger brother, as they experiment with various impromptu water sports. They finally settle on building their own little yellow pontoon boat in which they travel across the bay to school. Inspired by the reallife story of a yellow school boat in the Philippines, the linear narrative features ambient music and sound effects but no human voices.

Attack on Daddy is a production from the Korean Film Council. It features a combination of live action and animation. With realistic 3D effects, the user begins the experience inside the apartment home of a young Korean couple and their child, as seen and heard in a 360-degree 3 DOF format. The story soon becomes a fantasy adventure inside the little girl's dollhouse. Mostly told in the 3rd person, the story creatively and effectively incorporates two scenes in which the user enters a 1st person perspective. The linear narrative places the user in the role of father as villainous rabbits attack him inside the dollhouse, and then the user enters the point of view of a murderous 7' rabbit wielding a meat clever (combining animation and live action). Blending sounds effects, voices and music, the fast-paced experience is compelling.

Tale of the Tibetan Nomad is a UC Berkeley Theater Department production adapted from a play. The experience is a 360-degree live-action video with no

3D effects, although there are some interesting uses of depth of field. The linear narrative employs a dream sequence involving a Tibetan nomad who marries but loses his wife, only to awaken with renewed appreciation for the vows he is taking that day. Told using the 3rd person voice, narration (with limited dialog), acoustical effects, including ambient sounds, and traditional music advance the story. Audio cues encourage the user to look about and explore the visual scenes, particularly the colorful interior of the traditional tent home woven of yak-hair that the man occupies with his dream wife and newborn son.

Upstander is a 3 DOF 360-degree video animation from the U.S.A. It delivers an antibullying message, with scenes set inside a school gym, classroom, and hallway. Animated black and blue backpacks bully a little green one playing basketball, in the classroom, and at the hallway lockers, until a big red one intervenes. The narrative is in the 3rd person, linear and has no 3D effects, but incorporates some limited sound effects. Oculus
VR for Good supported the production.

Concluding Reflections

The entries in the Cinema360 competition provide a revealing look at the state of immersive cinema. There are promising productions from around the globe, and many of them are intriguing and compelling immersive experiences. Few of the productions, however, make highly effective use of the immersive platform's full potential. None of the productions feature 6 DOF; instead, they offer only 360-degree video and audio, no interactivity, and little use of NUI beyond simply looking about. AI and data use do not come actively into play, as there is no interactivity or haptic engagement, or customization or user control of the narrative flow. Several of the productions do utilize compelling 1st-person viewpoints, including *Saturnism* and *Home*, which employs simulated interaction by breaking the theatrical 4th wall. They are all relatively short in duration, though the experimental nature of cinematic VR of 2020 may account for that.

The 2020 productions are in contrast to *The Key*, a cinematic VR experience with interactive content and which won the Tribeca Film Festival 2019 award. *The Key* features an entirely 1st person experience that puts the user inside a mystery adventure classified as entertainment for persons ages 10+. *The Key* begins with Ana as host or friend who appears as alien and tells the user she has lost her memory and needs help. All she remembers is a mysterious key. The Oculus hand controllers become for the user mittens on their hands. When the user reaches out to grasp the key, the journey commences. It is worth noting that creating 6 DOF content requires different production processes than 3 DOF (Oculus, 2020), including stereoscopic, 3D cameras and for interactivity the use of an immersive game design engine such as Unity (SkarredGhost, 2020).

Limited to 3 DOF and featuring no haptic interface, Cinema360 entries underscore the potential for immersive experiences to be even more user engaging. In addition to enabling more active user mobility, designing narratives that utilize the Oculus haptic capacity could richly expand the multisensory user experience. Touching virtual objects and feeling them, tapping into proprioception, could greatly enhance the user's sense of self and presence in the virtual world. Sometimes called the sixth sense, or kinaesthesia, such haptic interfaces with proprioception could strengthen the user's sense of self-movement and body position within a virtual space and the apparent physical reality of that digital space. Of course, such capacity also demonstrates the importance of the Guardian virtual boundary. Whether the user is seated and stationary or standing and mobile in a room scale experience, physical movement while in a virtual environment carries with it the potential for physical injury as well as active engagement.

Cinematic VR that more fully utilizes the capacities of the VR platform may generate more powerful effects, including presence and empathy. Future research that focuses on these aspects of cinematic VR could provide valuable insight and guidance to media professionals and artists dedicated to creating new immersive experiences.

References

Archer, D. & Finger, K. (03/15/18). Walking in another's virtual shoes. *Columbia Journalism Review*. Retrieved 5 May 2020 from https://www.cjr.org/tow_center_reports/virtual-reality-news-empathy.php

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do.* W.W. Norton & Company: 1.

Baum, L. F. (1901). *The master key.* Bobbs-Merrill.

Collins, B. (4 May 2020). VR headsets are dying a lonely death. *Forbes*. Retrieved 4 May 2020 from https://www.forbes.com/sites/barrycollins/2020/05/04/vr-headsets-are-dying-alonely-death/

De la Peña, N. (2017). Greenland melting. *Emblematic Group.* Retrieved 5 May 2020 from http://emblematicgroup.com/experiences/greenland-melting/

De la Peña, N., Weil, P., Llobera, J., Giannopoulos, E., Pomés, A., Spanlang, B., Friedman, D., Sanchez-Vives, M.V., & Slater, M. (2010). Immersive

journalism: Immersive virtual reality for the first person experience of news. *Presence: Teleoperators and virtual environments 19*(4), 291 – 301.

Gilmour, R. (2012). 'Hunger in LA' makes its mark at Sundance.' USC. Retrieved 4 May 2020 from https://news.usc.edu/32639/hunger-in-l-a-makes-its-markat-sundance/

Gownder, J.P., Voce, C., Mai, M. & Lynch, D. (10 May 2016). Breakout vendors: Virtual and augmented reality. Retrieved 5 May 2020 from https://www.forrester.com/report/Breakout+Vendors+Virtual+And+Augmented+Reality/-/ERES134187

Hassenzahl, M. (2013). User experience and experience design. Chapter in A. Zahirovic, J. Lowgren, J. M. Carroll, M. Hassenzahl, & T. Ericsson (Eds.). *The Encyclopedia of Human-Computer Interaction (2nd Ed).* Interaction Design Fondation.

Höllerer, T., Feiner, S., & Pavlik, J. Situated documentaries: Embedding multimedia presentations in the real world. In: Proc. ISWC '99 (Third Int. Symp. On Wearable Computers), San Francisco, CA, October 18-19, 1999: 79-86.

Honan, M. (30 December 2013). I, Glasshole: My year with Google Glass. *Wired*. Retrieved 3 May 2020 from https://www.wired.com/2013/12/glasshole/

Jarrett, C. (28 November 2016). The psychology of eye contact, digested. *Research Digest.* The British Psychological Society. https://digest.bps.org.uk/2016/11/28/the-psychology-of-eye-contact-digested/

Kim, T. & Biocca, F. (09/97), Telepresence via delevision: Two Dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication.* Volume 3, Issue 2, page 0.

Lang, B. (20 March 2019). Oculus Rift S revealed with inside-out tracking, resolution bump, & new ergonomics. *Road to VR*. Retrieved 3 May 2020 from https://www.roadtovr.com/oculus-rift-s-specs-release-date-announcementgdc-2019/

Leprince-Ringuet, D. (April 28, 2020). AR and VR now have a chance to show what they can do. *ZDNet* Retrieved 28 April 2020 from https://www.zdnet.com/article/ar-and-vr-now-have-a-chance-to-show-what-the can-do-but-they-still-cant-break-through/

Mathur, M. B. & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition*, 146, 22–32. doi:10.1016/j.cognition.2015.09.008. PMID 26402646.

McLuhan, M. (1964/1994). *The Medium is the message*. MIT Press.

McQuail, D., Blumler, J.G., & Brown, J.R. (1972). *The television audience: A revised perspective. Media studies: A reader* (pp. 271-284). London: Longman.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Proceedings of Telemanipulator and Telepresence Technologies*, 2351–34.

Mittell, J. (2015). *Complex TV: The Poetics of Contemporary Television Storytelling*. NYU Press. ISBN 9780814769607.

Murray, J. H. (2017). *Hamlet on the holodeck: The Future of narrative in cyberspace*. The MIT Press. ISBN 978-0-262-63187-7.

Oculus (2020). Cinematic VR field-guide. Retrieved 5 May 2020 from https://creator.oculus.com/learn/cinematic-vr-field-guide/

Parisi, D., Paterson, M., & Archer, J. (2015). Call for publications: Haptic Media Studies,: *New Media & Society*. Retrieved 5 May 2020 from https://networks.h-net.org/node/7842/discussions/84847/cfp-haptic-media-studies

Pavlik, J. V. (2019). Experiential media and transforming storytelling: A theoretical analysis. *The Journal of Creative Industries and Cultural Studies, 3*, 46-67.

FIGURE 1: Narrative Qualities of 2020 Tribeca Film Festival Cinema360 Entries

Dreams to remember

1st Step: 360 imagery and sound with audio and visual directionality cues, 3D, 1st person, linear, documentary (Germany, 14:31)

Rain Fruits: 360 imagery and sound, mainly 3rd person, some 1st, linear, live action and animated (S. Korea, 14:03)

Dear Lizzie: 360 imagery and sound, mainly 3rd person, some 1st, linear, animated (USA, 2:45)

Forgotten Kiss: 360 imagery and sound with audio directionality cues, 3D, 3rd person, some 1st, linear, animated and live action (Finland, 13:24)

Seventeen Plus

A Safe Guide to Dying: 360 imagery and sound, 3rd person, some 1st, linear, limited nonlinear, live action (Greece, 12:41)

Black Bag: 360 imagery and sound, 3D, 3rd person, linear, animated (China, 12:36)

The Pantheon of Queer Mythology: 360 imagery and sound with audio and visual directionality cues, 3D, 1st person, linear, animated, interactive via web (Spain, 7:19)

Saturnism: 360 imagery and sound with audio and visual directionality cues, 3D, 1st person, linear, animated (France, 4:03)

Kinfolk

Ferenj: 360 imagery and sound, 3rd person, linear, live action (Ethiopia, 13)

The Inhabited House: 360 imagery and sound, 3rd person, linear, live action/archival photo/video (Argentina, 7:41)

Home: 360 imagery and sound with audio and visual directionality cues, 3rd person, linear, live action, simulated interaction by breaking 4th wall (Taiwan, 17:39)

Pure Imagination

Lutaw: 360 imagery and sound, 3rd person, 1st briefly, 3D, linear, animated (Philippines, 7:52)

Attack on Daddy: 360 imagery and sound with audio and visual directionality cues, 3D, 3rd person and 1st person, linear, live action and animated (S. Korea, 10:49)

Tale of the Tibetan Nomad: 360 imagery and sound, 3rd person, linear, live action (USA, 13:41)

Upstander: 360 imagery and sound, 3rd person, linear, animated (USA, 8:43)