Volumetric Video as a Novel Medium for Creative Storytelling

1

Gareth W. Young, Néill O'Dwyer, and Aljosa Smolic

V-SENSE, School of Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland.

ABSTRACT

Volumetric video (VV) is a media format that represents three-dimensional content for playback via traditional flat screens, 3D displays, and emergent extended-reality (XR) platforms, such as augmented and virtual reality (AR/VR). This type of volumography has numerous public-facing cultural applications within the creative technology domain and promises many advantages over conventional storytelling media. While many content creation approaches can call upon a long-established history of traditional storytelling techniques, the capture and creation of VV content and its representation with the XR space levels fundamentally new challenges that demand a redesigning of a storytelling grammar. This chapter discusses VV's application in creative experiments, explicitly outlining the V-SENSE project's innovative contributions to XR content creation. In this way, we explore VV as a novel medium for creative storytelling, providing example use-cases and real-world context for VV in practice. Our results demonstrate some successes and shortcomings of applications of VV in public-facing forums and reflective user-centered analyses. Ultimately, this chapter analyses the V-SENSE project's most complete VV productions and critically reflects upon case-study examples within the creative domain. Furthermore, it is expected that by reflecting upon the presented work, we can contribute to the establishment of a new XR storytelling grammar that will help provide solid foundations for other more sophisticated VV productions in the future.

KEYWORDS

Volumetric Video, Creative Experiments, Digital Arts, Practice Based Research

This chapter discusses the V-SENSE[1] project's contribution to volumetric video (VV) production as a novel medium for creative storytelling. As discussed in previous chapters, VV facilitates the 3D representation and visualization of dynamic, kinetic content (people, animals, objects, etc.) using live-action video instead of computer graphics and animation. VV can be recorded via several vision capture technologies, including infrared, RGB video, and LIDAR. For high-definition footage, dedicated studios have been created that surround the performance space with multiple cameras,

---

[1]https://v-sense.scss.tcd.ie/

FIGURE 1.1

Patrick Prendergast, the 44th Provost of Trinity College Dublin (2011-2021): (left) in the capture studio and (right) his keynote presentation via Microsoft's Hololens v1

capturing millions of data points for 3D digital reconstruction. Combined with 3D modeling techniques, VV contributes to imaginative story worlds for playback, broadcast, and display as dynamic 3D visual media that can be viewed across various platforms.

1.1 Volumetric Video and Creative Practice

VVs are different from traditional video media in two distinctive ways. Firstly, VVs are 3D and typically represented as 3D point clouds or 3D meshes (polygons). Secondly, VV offers 3D rendering capabilities with six-degree-of-freedom (6DoF), making them interactive, immersive, and expressive from a viewer-controlled first-person perspective. VV facilitates 6DoF by allowing audiences to freely change their viewing position (X, Y, Z) and orientation (pitch, yaw, roll). By contrast, conventional and modern 360° videos are presented on a two-dimensional (2D) plane or a three-degree-of-freedom (3DoF) sphere. Thus, classic 2D videos provide no viewport

FIGURE 1.2

Wireframe, solid, and textured 3D model (left-to-right)

freedom, and 360° videos only allow 3DoF (pitch, yaw, roll), with the viewer's normal, transverse, and longitudinal positions are permanently fixed or "baked in".

The VV pipeline is discussed in detail in other chapters of this book from capture to display. Despite significant progress, the technology still has limitations that may affect the visual quality and experience. These factors are related to the production costs of hardware (number of cameras, other sensors, studio equipment, etc.) and labor in post-production (interactive processing and refinement by experts using dedicated software tools). While high-end professional productions are expensive and well-equipped studios lead to professional quality results, lower-budget productions may compromise. Figure 1.1 illustrates an example of a person in V-SENSE's affordable VV capture studio (left) and the resulting VV viewed through AR glasses (right). Figure 2 illustrates the reconstruction process in post-production.

Maximizing quality and minimizing efforts in post-production, VV pre-production design typically considers some of the current limitations of the technology. This process includes the creation and use of costumes, accessories, hair, makeup, materials, etc. Certain colors may cause problems and reflective or transparent materials. Textured surfaces are well suited, while loose clothes may be problematic. Further, the specific movements of the performers may lead to inaccuracies, so choreography is also necessary concerning the speed and the nature of motion. Also, most studios are restricted by the number of individuals captured simultaneously (usually only 1 or 2 people concurrently). Despite these limitations, VV is increasingly used in multiple domains, with professional applications being developed for education, fashion, advertising, entertainment, and video games, indicating the stakeholders' interest in investing in the short-term with a view to the long-term gain.

VV can be regarded as a significant development in media production. Content for XR is becoming more pervasive in modern society for entertainment, healthcare, government, military, education, and industry training. With the rapid development and adoption of VR and AR technology in these areas, VV is also becoming a key technology for communicating meaning and artistic expression. Financial forecasts envision a positive future for the global VV market, estimating that it will grow from $1.4 billion in 2020 to $5.8 billion by 2025 [1]. This growth leads one to conclude that the future of VV for artists and producers is a bright one.

1.2   Case Studies

The V-SENSE project — Extending Visual Sensation through Image-Based Visual Computing (2016-2022) — was funded by the Science Foundation of Ireland (SFI) for research in creative technologies. The core technology enabling the creative productions presented in this chapter, i.e., VR/AR content creation based on 3D VV techniques, was developed by V-SENSE researchers [2] who spun out an innovative VV production company called Volograms[2].

During the life of the V-SENSE project, there have been numerous tests and

---

[2]https://www.volograms.com/

creative experiments, eight of which could be classified as significant standalone works of art: *MR Play* (2017–19), *The Trinity Library Long Room Mixed Reality Project, featuring Jonathan Swift* (2018), *Faoladh* (2018), *Bridging the Blue* (2019), *The Virtual Field Trip* (2021), and *Image Technology Echoes* (2020), as well as the forthcoming *Mixed-Reality Ulysses: Pilot Episode I* (2020) and *Episode II* (2022) and *XR Music Videos: featuring New Pagans* (2022). Six of these significant productions were made applying VV. What follows is a brief synopsis of these projects.

1.2.1   MR Play Trilogy (2017)

MR Play consists of a three-year trilogy of research experiments linking the performing arts and computer science disciplines [3]. Each part of the trilogy reimagines Samuel Beckett's groundbreaking theatrical text, *Play* (1963), for various formats of digital culture, including 1) a webcast (*Intermedial Play*), 2) virtual reality (*Virtual Play*) and, 3) augmented reality (*Augmented Play*). Play was chosen because it specifically engages the questions of dialogue and interactivity. In the original *Play*, the dramaturgy operates on the basis that the sequence of the actors speaking is determined by a moving spotlight, which Beckett likens to an 'inquisitor' [4, p.318]; they speak when the light is on them, and fall silent when the light is off. So, *Play* is a game of interaction between the light operator and the actor, mediated by light technology.

In the theatre, the audience passively observe the interaction between the light operator and the actors. The objective of the series of research experiments was to incrementally increase the audience's sense of immersion in the drama by exploring qualities such as agency and interactivity. In *Intermedial Play* the audience is drawn closer to the action through the employment of a pan-tilt-zoom (PTZ) robotic camera, where film grammar, such as super-close-up and sound effects, are used to create the illusion that the audience is complicit in the Pavlovian trail. *Intermedial Play* was the precursor, inspiration, and catalyst of *Virtual Play* and *Augmented Play*. *Intermedial Play* was a collaboration between Néill O'Dwyer (V-SENSE artist in residence) and Nicholas Johnson, Associate Professor in Trinity College's Department of Drama and secretary and co-director of the newly established Trinity Centre for Beckett Studies. An online video recording is available[3].

Virtual Play was a reinterpretation of *Play*, intending to engage a contemporary viewership via VR technologies (see Figure 1.3). In this version we acknowledge the interactive specificities of digital VR technologies and the role of the user as active; we recognise new opportunities for narrative and give the power of activation over to the end user, whose gaze becomes the spotlight. The user thus embodies the 'inquisitor' and is empowered to independently discover the story by looking at the characters and provoking them to speak.

This work was V-SENSE's inaugural arts and culture project under the creative technologies remit. The project was conceived to demonstrate how VR content could

---

[3]https://youtu.be/R11lT65HCZg

FIGURE 1.3

MR Play VV assets

be produced cheaply and expertly, thereby challenging the notion that sophisticated VR content is exclusively the domain of wealthy institutes and production houses. This virtual reality response to *Play* pushed the limits of possibility in consumable video and film by eliciting the new power of digital interactive technologies to respond to Samuel Beckett's deep engagement with the stage technologies of his day.

A central goal was to address ongoing concerns in the creative cultural sector regarding managing narrative progression in an immersive environment. The solution implemented here operated on the basis that users activated the virtual actors into speaking through their gaze, so the artwork acknowledged the new condition of active audiences and recognized new opportunities for narrative by affording audiences a central role in its unfurling. It was believed that by placing the audience at the center of the storytelling process, they were more appropriately assimilated to the virtual world and subsequently empowered to explore, discover, and decode the story instead of passively watching and listening. The gaming sector has effectively harnessed this narrative approach using procedural graphics and animation, but film and video (fundamentally based on live-action capture technologies) have struggled to engage this problem effectively. As such, this project investigated new narrative possibilities for interactive, immersive environments using live-action VV.

This VR version was, to our knowledge, the first full-length production of a VR drama using VV techniques. It is highly commended by both the computer science and digital arts communities, having been competitively selected for exhibition at some of the top art–science conferences and festivals in the world, including Beckett and Intermediality 2017 (where it premiered), New European Media 2017 (where it won first prize for the 'Art and Design Award'), SIGGRAPH 2018, Beyond Festival 2018 (at ZKM), the Prague Quadrennial 2019, and DRHA 2021 (where Professors Smolic and Johnson were invited to co-present a keynote on the subject). The work

FIGURE 1.4

Virtual Play (left) and *Intermedial Play* (right)

also produced a raft of publications on varying topics that intersect both the arts and computer sciences, including a description of the process [3]; two aesthetic reflections on practice that articulate the perspectives of the theatre practitioners and cogitate on the significance of the technological development [5]; a poster paper at a top computer science conference [6], which catalyzed a commission to expand the essay into an entire article for the world's leading art–science journal [7]; and, finally, a deep-dive human-computer interaction (HCI) study [8] that draws on the subjective opinions of industry and academic subject-matter experts and which teases out the potentialities, pitfalls, and areas for further research in this exciting, emerging genre of digital arts.

Augmented Play was the third and final part of the trilogy. The user and narrative paradigms were the same as the VR version; that is, users were invited to don the AR head-mounted display (HMD), embody the interrogator, and explore the narrative by confronting 'Beckett's characters virtually. This version had much in common with the earlier VR edition because it used the same VV assets, user interaction, and narrative development mode; however, the viewing paradigm was different because the content was displayed using cross-platform AR technology. Using AR allowed audiences to visually merge virtual, graphical, computer-rendered objects with real-world objects and scenes using a mobile phone, a tablet, or an HMD. Unlike VR, this approach allows users to see the world around them as it does not close off the outside world or fully immerse the audience in an 'other' computerized world. Therefore, the AR version elicited the exploratory specificities of the story by allowing people to interact with Beckett's characters in context-relevant, site-specific locations, meaning that it was ideal for location-based role-play and site-specific drama. As such, as a result

FIGURE 1.5

A visitor wearing the Hololens (left) and a screenshot of a user's point of view when wearing the Hololens (right)

of the after-life and sepulchral nuances of the story, *Augmented Play* was premiered in the vaulted, cavernous basement of the CHQ center in the historic Dublin Dock-lands Quarter to thematically engage the audience through contextualized, tangible architectural qualities to embellish notions of mystery, discovery, and reward.

1.2.2   Jonathan Swift at the Trinity Library Long Room

As well as holding a special place in the hearts of students, staff, alumni, and the wider Dublin community, the Trinity Library Long Room is a major destination for cultural heritage tourists. The objective of the Jonathan Swift project was to create an anecdotal XR narrative that would embellish and humanize the visitor experience by creating a welcoming and friendly digital character who is contextually relevant to the CH site and who adopts a 'humorous and playful mode of communication' via AR technology [9, p.22]. Because there is no substitute for the genuine experience of being physically present in the Long Room's breath-taking space, replacing the experience of visiting the library was not the initial goal of this project. Therefore, AR was deemed the most appropriate solution.

Specifically, visitors were to augment their real-world visit using an AR application that ran on a mobile phone, tablet, or an HMD device. These users were to be welcomed to the Long Room by a friendly VV representation of the famous Dean Jonathan Swift, who proceeds to divulge an embarrassing, whimsical and humorous memory of his youthful escapades while studying at the University. The application consists of a monologue about 1 minute in length. As in the previous *MR Play* project, both AR and VR prototypes were created using the same VV assets. Therefore, the XR technologies enabled an interactive narrative whereby a visitor could also remotely encounter the Swift character in a VR simulation of the Long Room.

The interactive VR prototype visualized the Long Room building and its contents, allowing the user to be immersed in the world of cultural heritage through simulation. The virtual architectural environment was built using a combination of manual 3D modeling and the computer vision process of photogrammetry, comprising a static 3D

model that could be imported into a game engine and then combined with the dynamic VV content. This approach meant that the visitors could be located anywhere on the planet and, by putting on an HMD, they could remotely enter the virtual Long Room and explore it via the various VR platforms available today.

The Trinity Long Room's Jonathan Swift XR project exemplifies V-SENSE's innovative approach to content creation, showcasing original technologies in real-world productions. The AR version occurs within the Long Room library's physical space using an AR HMD or a handheld mobile device (such as a smartphone or tablet). Using the spatial tracking features of AR devices, the digital volumetric video character appears standing beside a marble bust of Dean Swift (a permanent exhibit of the museum), as seen in Figure 1.5. The goal was to provide the various exhibitions with greater context and meaning, enhancing the overall visitor experience. This AR version was developed in the second phase of the project and only used the dynamic VV content; there was no need to virtually reconstruct the Long Room's internal, architectural spaces (the static content) because the user was already physically present in the geometric area. This production demonstrates VV content's versatility and illustrates novel tools and pipelines for creative digital content design.

The project's conceptual basis was founded on the idea that visualization technologies could create an engaging, interactive digital tour guide of the Long Room, featuring representations of famous historical personalities like Swift. The vision for the project was that the interactive tour would provide visitors with "peripheral stories that help draw their attention to the multitude of historical, architectural and archival details, procuring a deeper, more enriching experience of the world heritage site" [10]. The project was developed as a scalable pilot study focusing on a single example, with the view that it may one day be expanded to take in the entire museum by including multiple characters and personalities apt to the various exhibits [11]. Developing a commercial industry-ready AR tour guide application was beyond the means and remit of V-SENSE, but a quality of experience and grounded theory study was conducted, which presented, measured, and discussed "the appeal, interest, and ease of use of this ludic AR storytelling strategy mediated via AR technology in a cultural heritage context" [9]. The study's findings indicated "that humorous, playful storytelling is both appropriate and effective in this context because it enriched the visitors' experiences, and more of this type of work was encouraged" such projects, either in this museum or others.

1.2.3   Bridging the Blue

In this work, V-SENSE explored the role of immersive content creation in technologically mediated perspective-taking experiences of VR. *Bridging the Blue* was a creative experiment that explored VR as "the ultimate empathy machine" [12] where users could explore an imaginary world and experience personal depictions of clinical depression, experienced first hand by the artist. The experience harnesses VR as a new type of human-orientated technology that can be used for perspective-taking by applying a volumetric capture of real-world human performance to assess the effects

FIGURE 1.6

Bridging the Blue scene selection (Left) and narrative experiences (Right)

of empathy-building experiences achieved via new media [13]. Audience attitudes towards the protagonist were augmented to re-evaluate their experiences of perceived oneness and increase the amount of effort spent on helping oneself, others, and meaningfully engaging with society. From an industry perspective, this technological intervention was used to discuss production methodologies and interaction techniques when applied in XR content creation and to inform future creative technology projects.

The philosophies of augmented virtuality [14] were explored through the live-action capture and representation of the performer who was projected into and interacted with the virtual world; that is, objects from the physical world could co-exist with digital objects within the immersive virtual environment (IVE) (Figure 1.6). By merging the real – the volumetric capture of the narrator – and the virtual – the born-digital world designed and built using game engines and 3D modeling software – *Bridging the Blue* produced a VR platform to explore experiences of clinical depression.

Bridging the Blue was presented as an interactive VR experience that transported the viewer to a familiar yet imaginary world. The user begins on a small, rocky island out at sea (Figure 1.6), and they are surrounded by symbols that have significance to the artist and their experience of clinical depression. These artifacts act as interactive keys that allow the user to teleport to linked narratives (7 in total) that employ two main narrative methods: 1) the fly-on-the-wall method, where the (unseen) user observes a VV representation of the artist/subject in the middle of an episode without being able to interact with the scene, and 2) the first-person narrative method, where the artist/subject directly addresses the user in a sort of lecture format, describing the experiences and divulging information comprised of her ordeal and clinical facts. "Here, the technology has been harnessed as a device to catalyze shifts, prompting the viewer to evaluate and recalibrate modes of listening in the context of mental health" [15, p.17]. This goal is pursued through the organization of three main principles: 1) by assigning agency to the user, whereby they can navigate between narratives without the promise of rewards or an overall linear goal; 2) through heightened awareness and flow catalyzed by the novelty-factor of the technology; and 3) by simulating an embodied, experiential encounter with the artist that is pivotal to the transformative

intentions of the work. Interaction with the artist/subject is limited to listening. The intention is to model "responses more consistent with empathy, support, and validation in the context of non-professional conversations about mental health where the natural tendency is often to attempt to offer solutions or advice without having the requisite professional expertise" [15, p.18]. The project is an exercise in employing immersive technologies as a tool for subconsciously training best practice techniques, such as listening skills for people with no knowledge of mental health support protocols or skills in navigating such conversations.

Within each scene, the work challenges common misconceptions of and responses to depression and offers subjective playback and validation to those who have also been affected by this condition. In this context, the work demonstrates VR's new, innate capabilities and its capacity to inform the viewer. The artist [15] argues this in the context of the literary technique of defamiliarization, an aesthetic philosophy championed by the Russian Formalists, perhaps most notably by Viktor Shklovsky in his 1917 essay "Art as Technique" [16]. Given that Bridging the Blue is an immersive experience permitting an embodied, first-person encounter Arielle argues that under Held and Hein's technology–perception experiments [17], the venture enables a neurobiological basis for changes in perception. Human perception is a faculty of awareness that is malleable and subjective, and ordinarily, it is "habitual, economical, automatic—too familiar, in fact" [18, p.28]. Defamiliarization describes the literary technique of taking the audience out of their comfort zone by making ordinary and familiar objects appear different, causing them to question their preconceptions about the thing under regard, ultimately transforming perception. Arielle's strategy in Bridging the Blue is to challenge the audience's expectations of what they believe clinical depression to be by challenging them to witness symptoms and listen to its effects from a different perspective.

1.2.4 Image Technology Echoes

In the immersive experimental fiction *Image Technology Echoes*, the user enters a quiet gallery inside a cavernous museum space (Figure 1.7). *Image Technology Echoes* asks how we experience perception and embodiment and how much could be happening below the surface of those we meet. Real actors are captured using VV techniques and perform all of the characterizations and representations of the three scenes. The script for each set is written in collaboration with deep learning frameworks that use natural language processing (NLP) algorithms. The text generated by the algorithms was supervised and edited by the director/playwright Lauren Moffat and periodically reprocessed until the machine-human-machine partnership developed a suitably nuanced text that evoked the spirit of the artist's vision. In this regard, in terms of visual computing and the playwrighting process, *Image Technology Echoes* shows provocative new ways of creating work in collaboration with machines.

The narrative comprises three scenes: the gallery, which is constituted by a dialogue between the two main protagonists, and the two subconscious scenes formed by monologues. The gallery is empty, except for an older man and a younger woman,

FIGURE 1.7

Image Technology Echoes

who have a conversation about the painting they are looking at: a large expressionist canvas depicting a stormy ocean. The starched conversation circulates indefinitely unless the visitor enters the body of one of the two figures, at which time they are transported into the respective subconscious mental space of that character. These mental spaces manifest as messy rooms with a large window that overlooks a humongous version of the painting, which appears very different from each character's point of view. Each room contains a character that echoes the personality of its owner. The user has a fly-on-the-wall perspective in these rooms, watching the character's doppelganger or homunculus reciting a stream of consciousness. In a voyeuristic dynamic, the audience witnesses a personal, introspective, and slightly obsessive rant.

In September 2020, *Image Technology Echoes* was included in the competitive Vancouver International Film Festival as one of its Featured Projects – Volumetric Market. It was also displayed as a prototype demonstration at the Espronceda Institute for Arts and Culture in Barcelona in October 2020. *Image Technology Echoes* celebrated its world festival premiere at the International Film Festival Rotterdam in June 2021 and has been very well received across many of the top global VR communities since. In May 2021, *Image Technology Echoes* won first place at the prestigious VR ART PRIZE, awarded by Deutsche Kreditbank (DKB) in cooperation with the Contemporary Arts Alliance (CAA) Berlin featured in an exhibition at Haus am Lützowplatz April-June 2021. It also received acclaim during this time in the American Business Magazine, Forbes [19].

FIGURE 1.8

MR Ulysses

1.2.5   Mixed Reality Ulysses

In the *Mixed Reality Ulysses* XR application, users were invited to enter James Joyce's literary world Ulysses through AR and VR technology (Figure 1.8). The VR application allowed audiences from any part of the globe to experience the sites and associated scenes from the story via a VR headset. On the other hand, the AR application encouraged audiences to physically go to those locations and witness the VV as dramatic recreations of the scenes using mobile phones, tablets, or AR HMDs. At SIGGRAPH Asia 2021 [20], these experiences were fully prototyped as XR storytelling applications for a broader project that aims to depict multiple scenes of the book using VV capture techniques for the dynamic content and photogrammetric practices for the 3D scenic construction.

The pilot episode depicts a recreation of the book's opening scene, featuring a dialogue between Buck Mulligan and Stephen Dedalus (Joyce's pseudonymous doppelgänger) on the roof of the Martello Tower at Sandycove . The mode of audience interaction (in the pilot episode) consists of the user embodying Dedalus's character with Mulligan positioned opposite as the interlocutor. Therefore, Mulligan is the only visible character in the scene. The user is not required to articulate the words of Dedalus; these sections were delivered by the application and were heard (in full stereo) as if spoken in the first person. Mulligan's voice is perceived as coming from the other dialoguing character; his voice is subjected to 3D spatial audio — simulating distance and direction — to enhance immersion and embellish the sense of presence. The excerpt is a reinterpretation of the section where Stephen accuses Mulligan of speaking insultingly about his mother's death. Mulligan, a doctor by profession, scrambles to defend his position but his attempt to justify himself, in a speech about the normality of death, only causes further offense to his sensitive listener. While only three minutes long, the address and dialogue are profoundly

expressive and provide an insight into Joyce's dexterous handling of language.

Although hindered by the onset of the COVID-19 pandemic, the project was auspiciously timely because it provided insights into many of the vocational problems creative artists and performers encountered during the 2020/21 global lockdown. For example, in the AR instance, *MR Ulysses* demonstrates a site-specific theatrical paradigm where audiences could experience theatre and performance en plein air using mobile and wearable technologies. Furthermore, Bloomsday's annual literary festival, which since 1963 celebrates the book through site-specific reenacts of text passages, was canceled for the first time in 2020. In the VR instance, MR Ulysses provided a means for audiences to immerse themselves in a simulated spatial experience of Bloomsday without having to travel to Dublin physically by applying open-source photogrammetry processes [21,22]. *MR Ulysses* does not claim to be a replacement for the joy of site-specific physical gatherings that celebrate literature, but it does help provide an insight into how the vocational disenfranchisement (suffered by theatre/performing arts practitioners as a result of the global lockdown) can be eased by harnessing the global interconnectivity of electronic networks [23]. As such, *MR Ulysses* demonstrates how the paradigm of spatialized storytelling that is unique to MR can be effectively integrated with VV techniques in a way that challenges our current understanding of physical experiences of live theatre [23].

1.2.6 XR Music Videos

Finding new ways to visualize and express musical performance in VR is driven by artistic creativity, a desire to innovate technologically, and a need to capture new and existing audience attention. If "Artificial reality is the authentic postmodern condition, and virtual reality its definitive technological expression" [24, p.169], then postmodernist representations of art can be expressed from within VR. Immersive XR music videos were studied at V-SENSE as an emergent art form in and of themselves as many emergent XR technologies are being applied in this endeavor, such as stereoscopic and 360° audiovisual spatial recording technology. These capture technologies expanded the traditional viewing medium to include further dimensions of immersion, interaction, and imagination for the audience and were closely tied to advancements being made in home PC GPU/CPU speeds, HMD optics, software data processing capabilities, and AI.

Music videos are short films that integrate songs and imagery produced for artistic and promotional purposes. Modern music videos apply various media capture techniques and creative post-production technologies to provide a myriad of stimulating and artistic approaches to audience entertainment and engagement for viewing across multiple devices. Within this domain, VV capture technologies have become a popular means of recording and reproducing musical performances for new audiences to access via traditional 2D screens and emergent XR platforms, such as AR/VR. These 3D digital reproductions of live musical performances are captured and enhanced to deliver cutting-edge audiovisual entertainment. However, the precise impact of VV in music video entertainment was still in need of exploration from a user per-

FIGURE 1.9

Imaginative XR Music Video experiences can take you anywhere

spective. The XR Music Video project demonstrated how users responded to VV representations of music performance via XR technology. This approach explored how audiences were likely to react to music videos in an XR context and offered insights into how future music videos may be developed. The findings were a formative starting point for more sophisticated, interactive music videos that could be accessed and presented via emergent platforms, for example, how VV XR music videos may be captured, edited, and accessed for live performance.

The study of XR music videos was used to inform V-SENSE's user-centered design of a custom-made VV VR music video experience [25] , featuring the New Pagans' track *Lily Yeats*. The project's pilot study initially highlighted the specific qualities that audiences seek during the consumption of such materials. Iterations of this novel application area are expected to focus on differences between traditional media and new XR experiences and expose and build upon existing HCI studies that focus on music and technology in use, specifically those concerning how users experience music videos presented via 6DoF XR technologies.

1.3   Reflection on V-SENSE Practices

The aforementioned collection of V-SENSE's creative experiments has highlighted the potential of XR technology to shift artist and audience engagement paradigms. This context-specific focus spotlights how new uses for creative technology can bring further understanding and meaning to XR within the film, theatre, and performance domain. From exploring the innovative works of the V-SENSE project, we have observed a potential disruption to the creative cultural sector caused by the increasing

adoption of XR technologies in emerging modes of creative practice and evolving audience consumption behaviors. Following a series of user studies that use these experiments as stimuli, we can also understand and identify the explicit need to establish a new storytelling grammar [8].

A new storytelling grammar for XR must first balance the artist and audience perspectives and prioritize specialization and mise-en-scene over traditional temporal paradigms. XR must also facilitate entertainment by employing those specificities that come into their own in digital media: interaction, dialogue, exploration, discovery, networking, etc. Therefore further research is required to explore and define these qualities and reveal how these disruptive technologies will affect the creation and consumption of future performing arts practices.

This chapter highlights numerous considerations for the deployment of VV techniques in creative performance practice and the establishment of a storytelling grammar, surfacing in the cross-section of presented works and the qualitative and quantitative findings they produced. For example, careful consideration should be given to the role of temporality, spatiality, and the mixing of physical and digital realities when creating IVE experiences for VR and to site-specificity when using AR. However, these specificities of the XR platform also need to be meaningfully contextualized to the content to avoid the presented works being overwhelmed by novelty and technical exhibitionism. Existing performances will need to be re-established to engage audiences effectively. This factor mainly concerns timing and pace, interactive narrative models, and reconciling the shared audience experience.

In summary, it can be asserted that each VV-XR production engages the emerging genre uniquely and experimentally by tackling different aspects of the content-versus-technique relation, contributing to the definition of a nascent storyworlding grammar. The MR Play series engages audience interaction by placing the user at the center of the story's unfolding. The Jonathan Swift app for the Trinity Library explores the embellishment of the museological experience. Bridging the Blue investigates employing the technology to educate through a combination of listening and exploration. Images Technology Echoes explores the use of NLP for script development and also uses exploratory devices but, in this context, to challenge the subjectivities of the unconscious and human relations. In the context of intangible cultural heritage, XR Ulysses explores, on the one hand, geospatial AR technologies for site-specific performance and, on the other, VR technologies as a means of simulating site-specific performance. Finally, The XR Music Video investigates the potential of audiences as a creative agent by giving them control over some aspects of the musical tracks and the spatial relationship to the characters (aka the musician's avatars).

While our collection of experiments contributes to defining a new taxonomy and can help us begin to map the grammar of VV-XR, they by no means come close to covering the totality of the emerging genre. Just as filmmaking took half a century to solidify a widely accepted grammar and vocabulary, so will XR take a similarly long time for its language to be defined. In terms of history and the artistic avant-garde, what is certain is that with the emergence of new technologies, there follow a plethora of creative productions that attempt to harness them towards creative expression.

Just as technology is a continually evolving phenomenon, so too does art continually evolve; the grammar that emerges is determined by the gradual interplay of successes and failures that often swap position depending on the audience's subjectivity.

1.4 Conclusion

XR as a media entertainment platform can be much more than immersive gaming, just as film entertainment is more than blockbuster cinema. Contemporary XR productions offer an expansive entertainment value to existing and emergent 3D media practices and beyond. By approaching the use of technology creatively, objects of attention within an IVE can be presented so that audiences can move around, interact, and engage with creative materials, making the XR experience fundamentally unique and rewarding to the audience. However, this technological intervention can arguably only add further value over more traditional multimedia practices as a supplement to the main creative focus, such as music, cinema, gaming, and so forth, by making full use of the 6DoF experience as well as implementing engaging interactions and introducing player/audience agency. Although highly immersive and engaging, VV also brings light to issues concerning conventional performance practice in and for IVEs and how artists can audiences with artistic content remaining as intended.

Bibliography

[1] Market and Markets Research, Volumetric video market with covid-19 impact by volumetric capture (hardware, software, services), application (sports, events, and entertainment, medical, advertisement, and education), content delivery & region - global forecast to 2026 (2020).
URL https://www.marketsandmarkets.com/Market-Reports/volumetric-video-market-259585041.html

[2] R. Pagés, K. Amplianitis, D. Monaghan, J. Ondřej, A. Smolić, Affordable content creation for free-viewpoint video and vr/ar applications, Journal of Visual Communication and Image Representation 53 (2018) 192–201.

[3] N. O'Dwyer, N. Johnson, E. Bates, R. Pagés, J. Ondřej, K. Amplianitis, D. Monaghan, A. Smolić, Virtual play in free-viewpoint video: Reinterpreting samuel beckett for virtual reality, in: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), IEEE, 2017, pp. 262–267.

[4] S. Beckett, Play, in: The Complete Dramatic Works of Samuel Beckett, Faber & Faber, London, 2006, pp. 305–320.

[5] N. O'Dwyer, N. Johnson, Exploring volumetric video and narrative through samuel beckett's play, International Journal of Performance Arts and Digital Media 15 (1) (2019) 53–69.

[6] N. O'Dwyer, N. Johnson, R. Pagés, J. Ondřej, K. Amplianitis, E. Bates, D. Monaghan, A. Smolić, Beckett in VR: Exploring narrative using free viewpoint video, in: SIGGRAPH 2018 Posters, ACM, 2018, pp. 1–2.

[7] N. O'Dwyer, N. Johnson, E. Bates, R. Pagés, J. Ondřej, K. Amplianitis, D. Monaghan, A. Smolic, Samuel beckett in virtual reality: Exploring narrative using free viewpoint video, Leonardo 54 (2) (2021) 166–171.

[8] G. W. Young, N. O'Dwyer, N. Johnson, E. Zerman, A. Smolic, Mixed reality and volumetric video

in cultural heritage: Expert opinions on augmented and virtual reality, in: International Conference on Human-Computer Interaction, Springer, 2020, pp. 195–214.

[9] N. O'dwyer, E. Zerman, G. W. Young, A. Smolic, S. Dunne, H. Shenton, Volumetric video in augmented reality applications for museological narratives: A user study for the long room in the library of trinity college dublin, Journal on Computing and Cultural Heritage (JOCCH) 14 (2) (2021) 1–20.

[10] N. O'Dwyer, J. Ondřej, R. Pagés, K. Amplianitis, A. Smolić, Jonathan swift: augmented reality application for trinity library's long room, in: International Conference on Interactive Digital Storytelling, Springer, 2018, pp. 348–351.

[11] E. Zerman, N. O'Dwyer, G. W. Young, A. Smolic, A case study on the use of volumetric video in augmented reality for cultural heritage, in: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, 2020, pp. 1–5.

[12] C. Milk, Ted2015: How virtual reality can create the ultimate empathy machine (2015).
URL https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine?language=en

[13] G. W. Young, N. O'Dwyer, A. Smolic, Exploring virtual reality for quality immersive empathy building experiences, Behaviour & Information Technology (2021) 1–17.

[14] P. Milgram, F. Kishino, A taxonomy of mixed reality visual displays, IEICE TRANSACTIONS on Information and Systems 77 (12) (1994) 1321–1329.

[15] L. G. Arielle, Bridging the blue, in: R. Brown, B. Salisbury (Eds.), The Art Exhibit at ICIDS 2019 Art Book: The Expression of Emotion in Humans and Technology, Carnegie Mellon University, Pittsburgh, 2020, Ch. 1, pp. 15–28.

[16] L. Crawford, Viktor shklovskij: Différance in defamiliarization, Comparative Literature (1984) 209–219.

[17] R. Held, A. Hein, Movement-produced stimulation in the development of visually guided behavior., Journal of comparative and physiological psychology 56 (5) (1963) 872.

[18] D. P. Gunn, Making art strange: A commentary on defamiliarization, The Georgia Review 38 (1) (1984) 25–33.

[19] S. Rabimov, Why first virtual reality art prize is perfect for 2020 e (2020).
URL https://www.forbes.com/sites/stephanrabimov/2020/11/27/why-first-virtual-reality-art-prize-is-perfect-for-2020/

[20] N. O'Dwyer, G. W. Young, A. Smolic, M. Moynihan, P. O'Hanrahan, Mixed reality ulysses, in: SIGGRAPH Asia 2021 Art Gallery, SA '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 1. doi:10.1145/3476123.3487880.
URL https://doi.org/10.1145/3476123.3487880

[21] O. Dawkins, G. W. Young, Engaging place with mixed realities: Sharing multisensory experiences of place through community-generated digital content and multimodal interaction, in: International Conference on Human-Computer Interaction, Springer, 2020, pp. 199–218.

[22] O. Dawkins, G. W. Young, Workshop—ground truthing and virtual field trips, in: 2020 6th International Conference of the Immersive Learning Research Network (iLRN), IEEE, 2020, pp. 418–420.

[23] N. O'Dwyer, G. W. Young, A. Smolic, XR ulysses: Addressing the disappointment of canceled site-specific re-enactments of joycean literary, cultural heritage on bloomsday, International Journal of Performance Arts and Digital Media (2022) 1–19doi:https://doi.org/10.1080/14794713.2022.2031801.

[24] B. Woolley, Virtual worlds: A journey in hype and hyperreality, Benjamin Woolley, 1993.

[25] G. W. Young, N. O'Dwyer, M. Moynihan, A. Smolic, Audience experiences of a volumetric virtual reality music video, in: Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces, 2022, pp. 1–7.