Published by :

EARLY VISUAL MEDIA LAB
CICANT

EXPERIENCE DESIGN FOR VIRTUAL REALITY. FROM ILLUSION TO AGENCY

VOLKER KUCHELMEISTER
UNSW University of New South Wales, Sydney. Art & Design.
felt Experience and Empath Lab (fEEL - Australia)

Volker Kuchelmeister is an expert in presence, embodiment and place representation for immersive applications and has worked extensively in interactive narrative and immersive visualisation while exploring and exploding the boundaries of the cinematic image.

ORCID: 0000-0001-9161-788X

Corresponding Author:
Volker Kuchelmeister
UNSW University of New South Wales
Greens Rd, Paddington, NSW, 2021, Australia
kuchel@unsw.edu.au

Peer review:
Filipe Luz
CICANT/ Lusófona University (Portugal)
José Neves
CICANT/Lusófona University (Portugal)

INTERNATIONAL JOURNAL ON STEREO & IMMERSIVE MEDIA, Vol. 4 Issue no. 1
pp. 146-167
DOI: 10.24140/ijsim.v4.n1.08
ijsim.ulusofona.pt

Abstract

Virtual Reality (VR) allows viewers to inhabit and interact with virtual spaces in a way that has the potential to be much more compelling than any other medium, breaking through the barrier between merely watching to experiencing a situation or environment. It has an experiential quality by integrating the domains of interactive video games, filmmaking, storytelling and immersion. It is a balancing act between narrative design, digital placemaking and user agency. In this article, written from a practitioner's perspective, I propose and demonstrate strategies which allow immersive experiences to utilise multiple modes of representations, such as omnidirectional stereoscopic video and real-time 3D rendered geometry, to form a coherent spatial narrative environment for a viewer in VR. Particular emphasis will be placed on human factors, medium specific aspects and technological features. This insight emerged from a series of recent VR projects, which are fundamentally different in terms of content, design and production techniques, but this diversity is an opportunity to demonstrate how these factors can be applied across a broad range of projects.

Keywords: Virtual Reality, Immersion, Presence, Depth Perception, Spatial Narrative

Introduction

Over the last five years, a new wave of affordable and powerful VR hardware has sparked an incredibly broad range of applications spanning from entertainment across the arts to education and journalism, to name just a few. While consumer uptake is relatively slow, there is now a multitude of distribution channels for content, and VR is now firmly established in the cultural sector, and can be experienced by the public in museums, art galleries and film festivals. In research, the volume of publications related to VR has increased significantly, and new labs and conferences have emerged across the world. It is clear, there is a strong interest in this medium across many sectors and disciplines. Significant progress has been made in developing the technology, uncovering psychological and perceptive factors, and on the theoretical and philosophical front. What is not extensively studied is the diversity and multitude of VR experiences. In comparison, film-studies had over a century to analyse, categorise and better understand the language of film. With VR, we are just at the beginning, while trying to keep up with this rapidly evolving medium. The study of experience design for VR needs to consider the complexities of the medium, while analysing the form and the genre of an experience, the relation to other media-studies and how it is perceived by the viewer. This article is not intended to be a comprehensive study in immersive experience design; this has been expertly achieved by Sherman et al. (2002, 2018) and others. It provides insight from a practitioner by first defining the principal human as well as technological medium specific factors, and secondly describing how they were applied to a number of selected projects.

Immersion and Presence

A central property of a compelling VR experience is the sense of presence a viewer might be able to experience, afforded by the immersive technology. Immersion is the objective degree to which an immersive system and application projects stimuli for the sensory receptors of users in a way that is extensive, matching, surrounding, vivid, interactive, and plot informing (Slater and Wilbur, 1997). Therefore, immersion is the property of the medium; it presents stimuli which are perceived and interpreted by an individual. How the viewer subjectively experiences the immersion is known as presence. With presence, a viewer might feel more like visiting a place than merely perceiving it. Or as defined by Lee (2004, p. 32), presence is ‹the psychological state in which the virtuality of experience is unnoticed' or in Davis' words (1998, p. 247),, 'The concept is absolute simulation: a medium so powerful that it transcends mediation, building worlds that can stand on their own two feet.'

However, immersion does not always lead to a sense of presence. On a technical level, it is necessary for the immersive system and application to afford agency and simulate human perception. The level of presence is not only related to the number and fidelity of perceptual cues (the realism), or the degree of agency for a participant. As a product of the mind, it is also described as a *feeling* of 'being there'. Diemer et al. (2015, p. 6) have shown that emotions, as much as the immersive quality, positively influence the level of presence experienced: 'the stronger the feelings involved, … , the greater the likelihood of finding a significant correlation between presence and emotion'. This is a significant finding, indicating that presence is also a consequence of an emotional engagement, and vice versa, as Diemer et al. (2015, p. 4) observed: 'Presence is commonly regarded as a necessary mediator that allows real emotions to be activated by a virtual environment'. However, I believe more research is needed to determine to what degree the sense of presence or the subject of the experience is the primary factor for a strong emotional response. A storyteller or filmmaker might argue it is the narrative, visual language and imagination of the viewer that triggers the emotions.

Nevertheless, it is clear that presence is one of the primary factors for a VR application to produce a genuine and impactful experience. The perceptual illusion, combined with viewer agency and emotional engagement, has the potential to break the boundary between a mediated and a real experience. It may form an authentic episodic memory and be remembered later as a non-mediated event.

Beyond the Frame

Traditional visual media are fundamentally tied to a frame. It provides a compositional structure in the visual arts, photography and cinematography but also a connection to the world outside the image space and a frame of reference. In film and television, the viewer is always looking through the eyes of the camera into the narrative world that unfolds within the frame. Even as the image in the window is moving, it is the director and cinematographer who control the movement on screen, while the audience is restricted to a passive role. Friedberg (2006, p. 18) described '… the paradox of the mobility of the image versus the immobility of the spectator'.

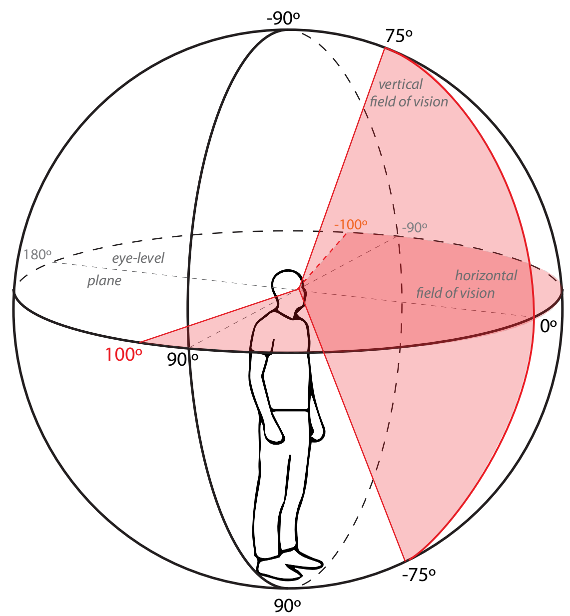

Fig. 1 The human field of vision, perspective view, including peripheral vision.

An omnidirectional image viewed within an immersive system, on the other hand, allows the spectator to control the direction of view and it eradicates the frame by a wide field of view and by blocking out any external visual stimuli. As Jones (2018, p. 39) noted 'we are provided an extensive space that is subject to our wandering gaze.' The image space becomes all-embracing and singular and the observer is completely subjected to the depiction. Filmmaker Alejandro G. Iñárritu noted in an interview relating to his cinematic VR film *Carne y Arena*, in which the viewer embodies a refugee, 'The basic idea is to experiment with VR in order to explore the human condition in an attempt to break the dictatorship of the frame, within which things are just observed, and claim the space to allow the visitor to go through a direct experience walking in the immigrants' feet, under their skin, and into their hearts' (Iñárritu, 2017, p. 1).

Going beyond the constraints of the frame requires spatial coherence between the viewer's gaze direction and the portion of the omnidirectional image presented by the system. Furthermore, it is essential to cover the human field of vision, including peripheral vision (Fig 1). This ultra-wide field of view as a continuous image space goes beyond what can be represented in linear perspective. Spherical 360º x 180º equirectangular or VR180 ≤180º circular fisheye projection expand on the limitations of linear perspective (Fig. 2).

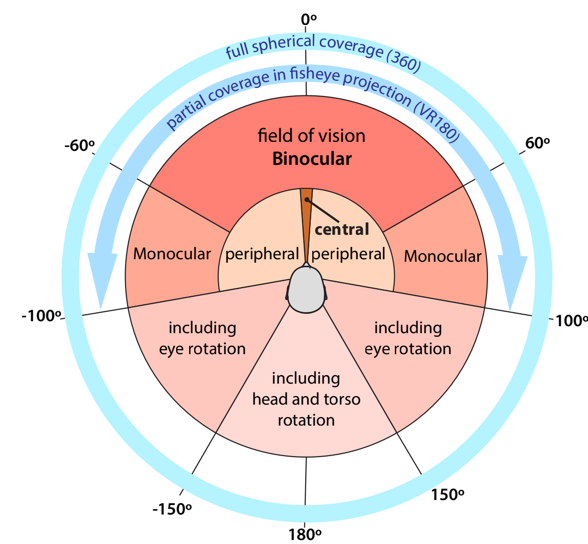

Fig. 2 Horizontal field of vision, also indicating binocular region, ego-motion and field of view for spherical and VR180

Depth Perception

Visual depth perception is the ability to perceive the world in three dimensions and judge the distance of objects. It arises from a variety of visual cues, classified as Monocular, Binocular and Proprioceptive depth cues (Fig. 3). Monocular depth cues are either pictorial, for instance perspective foreshortening, or rely on ego-motion to induce motion-parallax. Binocular depth perception is based on the principle of binocular vision, and the cognitive process of stereopsis, in which the brain estimates depth within the visual field based on the disparity of two slightly dissimilar images presented to the visual cortexes of both eyes. Various cues operate at different distances (Cutting and Vishton, 1995); for instance, binocular disparity is only effective within a range of ~20m (Fig. 4), and pictorial cues do not have a range limit. Stereoscopy is concerned with simulating binocular vision by technical means, be it with a Wheatstone Stereoscope (1830s), an apparatus in which the viewer can experience depth from a stereoscopic photographic pair, 3D cinema or a Virtual Reality headset. They are all based on the principle of channeling one view of a scene to one eye only and a

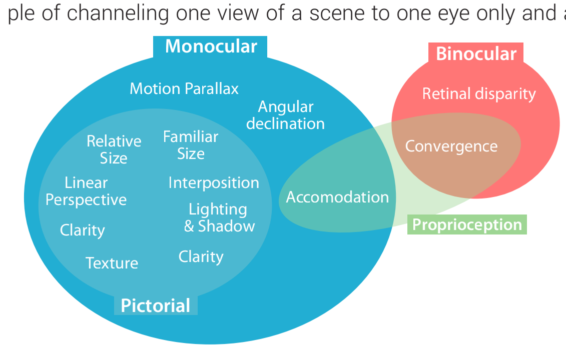

Fig 3 Depth perception cues. Classified as Monocular, Binocular and Proprioceptive.

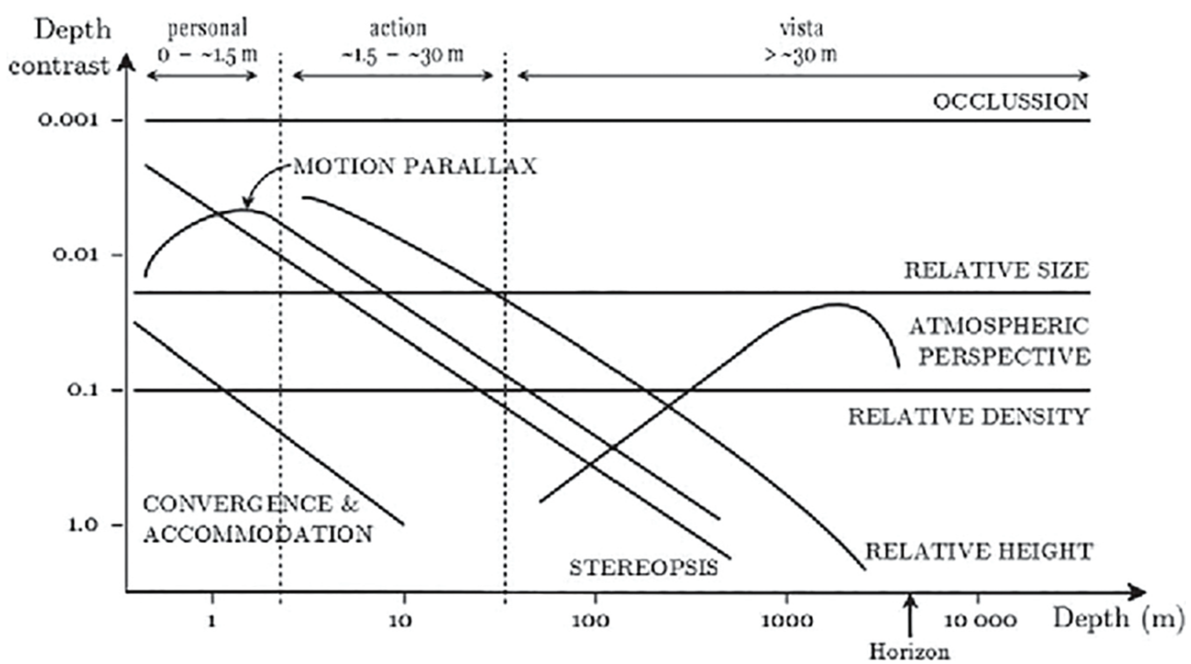

Fig. 4 The effectiveness of monoscopic and stereoscopic depth perception cues across a range of concentric regions: personal, action, and vista space (based on Cutting and Vishton, 1995,

disparate view to the other. In its simplest form, these stereoscopic image pairs are captured with two cameras spaced horizontally at a human interpupillary distance. However, there is a lot more to stereoscopic imaging; it is as much a technical challenge as an art-form. With various stereoscopic capture and display options, a scene or object might be perceived as tiny or gigantic, shallow or voluminous, distant or close. The image is transformed into a volume and the screen frame into a window. Stereoscopy is also an important tool to spatialise content. Objects and scenes within a stereoscopic representation are able to operate in negative parallax or in-front of the screen. The screen or projection plane forms the link between the virtual, stereoscopic world and the environment in which the images are viewed. Objects share the space with the audience and can produce a more tangible and embodied experience.

VR technology, with its inherent binocular design is an ideal platform to present stereoscopic 3D content. However, there is a fundamental difference between stereoscopic representation on a screen vs VR. In modern VR systems and omnidirectional image representation, there is no apparent

screen surface or frame. The notion of negative, zero and positive parallax and its spatial relationship to the screen surface and frame, important considerations in traditional stereoscopic imaging, are no longer relevant in VR. As a consequence, a wide field-of-view stereoscopic image viewed in VR loses its frame of reference to the world outside the stereoscopic image space. For VR, stereoscopic image pairs should be horizontally aligned in a way that stereo infinity; that is, the far plane, for instance a horizon, is at zero-parallax. Closer elements in the scene operate in negative parallax space. For a viewer, the optical characteristics of the VR headset push the scene across a wide range of depth.

Motion and the Space Around Us

The space around the body and the proximity of people or objects is of fundamental importance in terms of how we perceive, process and interact with the world. There are two disciplinary traditions in dividing this space. In neuropsychology, it is referred to as Peripersonal and Extrapersonal space (Rizzolatti et al., 1997) and the segmentation in perception research and social psychology is personal, action and vista space (Cutting and Vishton, 1995) (Fig. 5).

Peripersonal space is defined as the space immediately surrounding our bodies. Objects within can be grasped and manipulated. Objects located beyond in the extrapersonal space cannot be reached without moving towards them, or them moving towards us. Objects close to or in touch with our body are perceived with a multitude of senses and elicit enhanced neural and behavioral responses (Buffacci et al., 2018). We might therefore assume that the brain's processing of objects in peripersonal space is more thorough, more complex, and involves more modalities of sensory information than for objects located in extrapersonal space (Brozzoli et al., 2012).

The three concentric regions in perception and social science are subdivided into: personal, action, and vista space. Personal space is equivalent to peripersonal space, the volume within arm's reach. Within, we are typically only comfortable with others in intimate situations or due to public necessity (for instance in a busy subway). Working within personal space offers advantages of proprioceptive cues, more direct mapping between hand motion and object motion, stronger stereopsis and head parallax cues, and finer angular precision for motion (Mine and Brooks, 1997). Beyond this, action space is the region of public interaction. We can move relatively quickly within and speak to others. This space extends to about 20m from the viewer. Beyond the action space extends the vista space, where we have little immediate control and perceptual

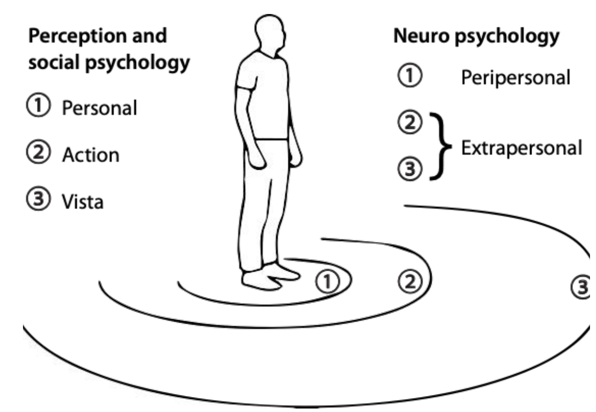

Fig. 5 The segmentation of the space around us as described by Cutting and Vishton (1995) and Rizzolatti et al. (1997).

cues are limited. Binocular vision is almost non-existent and depth cues are predominantly pictorial (Fig. 4).

However, space is dynamic and the world around us is never really motionless. Additionally, humans, and probably everything living, is hardly ever stationary. Observers move about in the environment, and observation is typically made from a moving position (Gibson, 1997). This subject-relative or ego-motion occurs when the spatial relationship between the observer and observed changes. Humans are quite sensitive to this type of motion, on a perceptive level, through proprioceptive cues and vestibular input. Motion induced parallax is one of the central factors in depth perception and plays a vital role in reconstructing and understanding the architecture of a place. Motion parallax relies on the spatial separation of a foreground object against a backdrop. A two-dimensional photograph does not create motion parallax, but the same framed image layered on a background will.

Relating this to VR, the design of a virtual world requires close attention and is of utmost importance. The placement of objects at a specific distance, their scale, movement and the ego-motion of the viewer have fundamental consequences with regard to cognitive processing as well as perceptual and social implications. For instance, objects that are close and move rapidly towards us might be perceived as a threat and viewers might feel uncomfortable with a character entering their personal space. However, as a film maker, I would not want to operate exclusively in the comfort zone. Just as in cinema, the close-up, for instance, allows a director to reveal previously unseen detail and more importantly, establish a strong emotional connection with the audience and create a sense of intimacy. The same is true for VR. In order to create an impactful experience, the director should work across the entire range of the time-space continuum.

This is, of course, not novel or specific to Virtual Reality. Visual artists, filmmakers, scenographers and so on have always been aware of, and worked with these principles. But it is valuable to contextualise VR experiences in this way and adapt these concepts with medium-specific factors in mind.

Cinematic vs Real-time VR

Cinematic VR, in the form of spherical video in omnistereo or monoscopic format generate a high level of realism and fidelity for a VR experience. However, the lack of positional tracking, and therefore the deficit in agency for a viewer, has a limiting effect on the perceived sense of presence in a scene. Real-time rendered 3D environments overcome this restriction; a viewer is able to explore freely, but often at the expense of realism.

In its simplest form, an immersive VR experience can consist of nothing other than a single monoscopic 360º photograph wrapped around the inside of a sphere. As we have seen, without any further spatial or motion cues, it is quite unlikely that this scenario will be able to produce a strong sense of presence for the viewer. Only with a range of cues can VR break through the barrier between merely watching to experiencing. Even by adding only binocular depth (presenting a stereoscopic 360º image), the sense of presence will greatly improve. However, as soon as the viewer physically changes position, the illusion of three-dimensional space diminishes. The viewer

has no agency, but can look around. To overcome this shortfall, the system needs to be able to respond to ego-motion by tracking the viewers' position in physical space and subsequent translating this information into the virtual environment. Real-time VR allows a viewer to freely explore the 3D world whereas in cinematic VR, the viewer is restricted to a predominately passive role (Table 1). This is also referred to as the degree of freedom (DOF), the number of axes an object can translate in three-dimensional space. There are a total of six degrees of freedom which describe every possible position and orientation of an object (left-right, up-down, forward-backward, pitch, yaw, roll). VR headsets and input devices are generally classified as 3DOF or 6DOF. They either track a viewer's head rotation only, or rotation and translation.

Cinematic VR, being a recorded medium, has an inherent realistic quality, whereas real-time rendered 3D content appears often stylised. 3D content and realism are directly proportional to the production value. The creation of a realistic real-time 3D scene is resource intensive, often limited to AAA computer games. One strategy to create a believable and realistic VR environment, while maintaining viewer agency lies in the combination and integration of both representational modes – a union between 3DOF and 6DOF to create a single coherent spatial environment. This hybrid approach is not yet widely adopted or studied, but I believe is an important step towards a fully volumetric representation in the future.

On a side note, the combination of immersive backdrop and three-dimensional foreground objects has a long tradition and was established with the panorama spectacle in the 19th century. A faux terrain (Fig. 6) is added to the foreground of the rotunda. This allowed for motion-parallax and provided a sense of depth to the otherwise two-dimensional panorama.

	Cinematic VR - 3DOF	Interactive real-time VR - 6DOF
Head tracking	Head orientation only. Stationary position.	Orientation and position.
User agency	Immotive. Ego-motion only between distinct locations. Optional timing control.3DOF controller.	Unrestricted continuous ego-motion within the room-scale tracking volume. Teleportation outside. Manipulation of objects with 6DOF controller or hand-tracking.
Media	360º or 180º video/photo or pre-rendered CGI scenes.	3D models, 3D scenes, textures, lights.
Format	Equirectangular format. Optional in stereoscopic 3D.	Real-time rendering in VR headsetspecific format. Stereoscopic 3D.
Binocular depth	Horizontal plane only.	Horizontal and vertical plane.
Audio	Pre-recorded and mixed. Optional Ambisonic or HRTF.	Real-time generated location sensitive. Optional Ambisonic or HRTF.

Table 1 Cinematic VR with three degrees of freedom (3DOF) compared to real-time VR systems and applications (6DOF).

Conclusion, Discussion and Appendices

In summary, this article has presented the fundamental building blocks related to experience design for VR applications, covering human factors, medium-specific aspects and technological features, to lay the foundation for the analysis of an extensive catalogue of VR productions in the appendices of this article. The aim is to demonstrate how these theoretical considerations are applied across a broad range of projects and to discuss strategies and techniques that can contribute valuable insight for designers and producers to create more effective and impactful immersive experiences. Furthermore, this practice-based research perspective might be a useful resource for media studies, investigating the production and content of this exciting medium.

However, theoretical reflections, descriptions, photos and video documentation can only go so far in documenting an immersive experience. The embodied and kinaesthetic aspects, the emotional engagement and the dimensional extent, central to being immersed and to feeling present in VR, is mostly lost in translation. I believe it is equally important to develop strategies to preserve those productions in their original format. With the rapidly evolving technology, whole systems become obsolete, and platform-specific content is no longer accessible. This has been the case since VR emerged in the 1990s, for instance, the VR artwork *Osmose* (1995) by Char Davis. Created with Softimage tools and developed for Silicon Graphics computers, it is now only available as a video recording. The same is true for a whole range of applications developed for Google Cardboard VR, a platform no longer supported.

Fig. 6 Painted panorama with faux terrain in Broken Hill, Australia. The Big Picture at Silver City Mint and Art Centre. A spectacle painted by Ando. The canvas measures 100x12m.

Following this conclusion, the appendices describe and partly dissect and analyse a number of VR projects that fall within the principles laid out in this article. Particular emphasis will be placed on factors in visual perception; experience design – including narration, scenography and user agency; and the technical implementation. The projects are classified across nine categories and grouped into Embodiment & Proximity; Empathy & Perspective taking; and Exploration, Play & Co-Presence.

More detailed information, full credits and video documentation can be accessed online. The URLs are listed in the project reference section at the end of this article.

Appendix A:
VR productions – Embodiment & Proximity

waumananyi: the song on the wind

2019, narrative 12:46, coda 10:10, language: Pitjantjatjara/ English

Narrative: *linear*; Point-of-view: *observer*; Agency: *explore (6DOF)*; Cast: *virtual human*;Production: *videographic*; *Stereoscopic*; Field-of-view: *360*; VR Platform: *6DOF*; Audio: *localised/mono*

This project explores mental health and healing from an Aboriginal perspective. Created by the acclaimed Uti Kulintjaku, formed from the Ngangkari traditional healers and artists, the collective addresses community issues of mental health from both Aboriginal and Western perspectives. *Waumananyi* is an Anangu (western desert people) led response to the experiences of constraint, entrapment and depression through the telling of the traditional story *The Man in the Log*. The virtual reality experience asks what it is like to be physically and mentally trapped in a space that you can't escape from. In this case the viewer embodies a hunter trapped in a hollow log. He gets glimpses but can no longer connect with people in the community. It is a metaphor for incarceration, separation, addiction and that sense of powerlessness that many people in aboriginal communities experience.

The story is told from an embodied perspective of the hunter (Fig. 8), and it transitions from the open desert landscape into the tight claustrophobic space inside the hollow log and back out into the open. This abrupt change from vista-space-dominated environment to the close proximity of the hollow log, intruding into the viewer's (peri)personal space, is intended to leave the viewer with a feeling of sudden entrapment.

The first-person perspective of the hunter was filmed stereoscopically with a VR180 camera, mounted close to the actor's eyes. The scenes within the virtual log were shot with a

Fig. 7 Cast: Pantjiti Imitjala Lewis, Rene Wanuny Kulitja, David Miller posing with stereoscopic 360° production camera. *waumananyi: the song on the wind* (2019). Photo by Rhett Hammerton.

stereoscopic 360° camera to allow the viewer to get glimpses in all directions and follow the narrative.

The hollow log is constructed as a 3D model from various pieces of bark (Fig. 9), photographed on site, processed with structure from motion and assembled in post production. As a 3D model, the log allows for full 6DOF freedom of movement, while the 360° stereoscopic video skybox background is stationary. Motion-parallax induced by head movement causes a shift in the perspective for the viewer, while looking through the cracks in the log. The holes reveal just enough to get fleeting glimpses of the protagonists and landscape outside. Over time, with successive attempts to free the hunter, the log becomes more porous.

As the story progresses, and day follows night and goes back to daylight; the virtual log is lit accordingly in the game-engine (Unity 3D), while the background changes to a night or daytime scene. This synchronisation of the mood in the scenes across the foreground log and background video helps to visually fuse the layers. The project exemplifies a hybrid approach, incorporating cinematic 3DOF and 6DOF elements.

Fig. 8 Embodied perspective: the hunter is freed and taken care of by his community. VR180 stereoscopic format. *waumananyi: the song on the wind* (2019).

Fig. 9 Point-of-view perspective inside the hollow log (top-left). External view with spherical image backdrop and 3D model of the log (bottom-left). Exhibition view (right). *waumananyi: the song on the wind* (2019).

At a public exhibition of the work, we asked the visitors to step into a physical representation of the hollow log, to create a direct relationship between the physical space and the VR environment. Thus, there was a visuotactile barrier to increase the sense of presence and to provide haptic feedback (Fig. 9, right).

Through the Eyes of the Mawa

2020, 6min, language: English. Proof of concept.

Narrative: *linear;* Point-of-view: *embodied;* Agency: *explore (6DOF);* Cast: *videographic;* Production: *videographic, 3D geometry;* Stereoscopic: *stereoscopic*; Field-of-view: *VR180;* VR Platform: *6DOF*; Audio: *ambisonic/ hrtf*

Initiated by Andrew Belletty, this project is a contemporary take on an old story from the Torres Straight island of Saibai. Artist, storyteller, choreographer and mask-maker Jeffrey Aniba Waia tells the story of a mask with significant cultural value, which was taken from the island by an early explorer and anthropologist, to end up in the British Museum. The narrative spans time and culminates in the return of the artefact to its origin. This narrative frames a more contemporary message about pressing problems the islander community is facing today, a rising sea level, plastic pollution and dwindling fish stock.

As the title suggests, the viewer embodies the mask character, following its journey through the centuries. The backdrop is captured with a VR180 camera in stereoscopic 3D. The

Fig. 10 *Through the Eyes of the Mawa* (2020). Composite image with backdrop, a placeholder 3D model for the mask and VR viewer.

Fig. 11 *Through the Eyes of the Mawa* (2020). Video still with Jeffrey Aniba Waia

mask is embedded in very close proximity to the VR viewer as a 3D model (Fig. 10-11). The model of the mask is only loosely attached to the VR head position and orientation; it seems to have its own will. This proximity goes beyond the viewer's personal space and feels almost like it is touching the skin, while the viewer feels the physical sensation of the head-mounted device. Furthermore, the mask forms a frame of reference: it partially obstructs the view and by being so close, it provides a very strong sense of depth through motion-parallax against, and in combination with, the stereoscopic video backdrop. To uphold the illusion of a coherent stereoscopic space between the video background and independent mask foreground, the perceived depth within the video does not intrude into the mask space. They are organised as separate layers, while at the same time forming a unified 3D space.

Appendix B: VR productions – Empathy & Perspective Taking

Being Debra

2019, 12:05, language: English.

Point-of-view: *embodied;* Agency: *observe (3DOF);* Cast: *videographic;* Production: *videographic, 3D geometry;* Stereoscopic: *stereoscopic;* Field-of-view: *VR180;* VR Platform: *3DOF;* Audio: *localised/ mono.*

Fig. 12 *Being Debra* (2019). Embodied perspective. On location.

Fig. 13 *Being Debra* (2019). A flashback scene.

Over 70 years ago, psychologist Alfred Adler described empathy as 'seeing with the eyes of another, listening with the ears of another, and feeling with the heart of another'. This project explores how technology can help promote empathy by affording the viewer an embodied experience of disability and its public reception. It investigates the influence affect and emotions have on presence and the potential of the medium to promote social change. *Being Debra* offers the audience a taste of the experience of being a dwarf in contemporary Australian society. The subject of this immersive experience is artist Debra Keenahan, who lives with achondroplasia, dwarfism. It includes flashbacks to Debra's memories growing up, as well as her routine daily encounters (Fig. 12-13).

To let the viewer experience the world through Debra's eyes, a stereoscopic VR180 camera is placed close to the first-person perspective of the subject (Fig. 14, left). This technique is effective for recreating the experience of unwanted intrusion into one's personal space, which occurs frequently in Debra's life. Cinematic VR experiences filmed in this way are inherently 3DOF and directional. They do not support ego-motion nor do they allow the viewer to glance over his shoulder. However, the wide field-of-view (≤180° horizontal and vertical) covers the entire field of vision, and in combination with binocular depth, provides an acceptable compromise. For this particular project, with the action staged in the centre, there is less motivation for the viewer to look back. This is really the case

Fig 14 Embodied first-person perspective. Stereoscopic fisheye camera system VR180 (left). Chroma key stage (right). *Being Debra* (2019).

for most situations; how often does someone turn around and look behind, if there is no good reason to do so?

The flashback scenes are situated within abstract representations of places, created in the VR drawing app Google Tilt Brush and rendered out as stereoscopic 360º still images. The flashback performances were shot in a chroma key studio (Fig. 14, right) and composited stereoscopically with the virtual set in post-production (Fig. 13).

The Quiet Discomfort

2018, 12 min, language: English. Proof of concept.

Narrative: *linear;* Point-of-view: *observer;* Agency: *explore (6DOF);* Cast: *videographic;* Production: *videographic, 3D geometry;* Stereoscopic: *stereoscopic;* Field-of-view: *360º spherical;* VR Platform: *6DOF;* Audio: *localised/ mono.*

Directed and written by Rosie Dennis (Urban Theatre Project, Sydney), *The Quiet Discomfort* grapples with the complex social issue of emotional violence towards women and children.

Fig. 15 Chroma key studio, utilising a stereoscopic mirror rig to capture video and a time-of-flight 3D sensor. The TOF data has been discarded for this production. *The Quiet Discomfort* (2018).

In this project, the viewer takes on the role of an observer, witnessing various domestic situations and narrative vignettes, played out in a home. It combines live actors, shot stereoscopically in a chroma key studio (Fig. 15) and a virtual set. The set, a domestic kitchen, is reconstructed with structure form motion as a realistic rendered 3D model (Fig. 16). The goal is to construct a coherent 6DOF experience for the viewer while maintaining a realistic representation of the cast and the environment and shifting the viewer's role from passive observer towards an active participant in the narrative.

On a technical level, the keyed stereoscopic video of the characters' performance is mapped onto two-dimensional sprites, which are placed in the virtual set (Fig. 16). In this way, the viewer has agency to move about while the narrative plays out, at least within a limited radius. The limit stems from the fact that the characters are only flat sprites within the 3D scene. They are rendered stereoscopically, but do not feature a three-dimensional volume as a fully digital character would. But within this limit, the ego-motion induced motion-parallax is a very effective cue for the viewer, and this strategy overcomes some of the limitations a 3DOF experience exhibits.

Fig. 16 The virtual set, a domestic kitchen, with embedded characters as stereoscopic video sprites (left). Viewer perspective (right). *The Quiet Discomfort* (2018).

Appendix C: VR productions – Exploration, Play and Co-Presence

EmbodiMap

2020, duration variable.

Narrative: *non-linear;* Point-of-view: *observer, embodied;* Agency: *explore (6DOF);* Cast: *virtual human;* Production: *3D geometry;* Stereoscopic: *stereoscopic;* Field-of-view: *360º spherical;* VR Platform: *6DOF;* Audio: *localised/ mono.*

EmbodiMap, developed at the UNSW felt Experience and Empathy Lab (feel), is a creative research and therapeutic tool that enables users to connect with and explore how thoughts, sensations and emotions are experienced in the body. It extends existing body-mapping research and protocols (De Jager et al., 2019) by facilitating a tangible immersive experience.

It consists of tools, protocols and facilitated engagements, that use 3D immersive drawing technologies (Fig. 17). Each time a sensation or emotion is identified, it can be mapped, by drawing into a figure. The participants can move around the immersive space to observe or enter into their mapping. *EmbodiMap* is designed to support a single user or a remote session with multiple participants represented as avatars

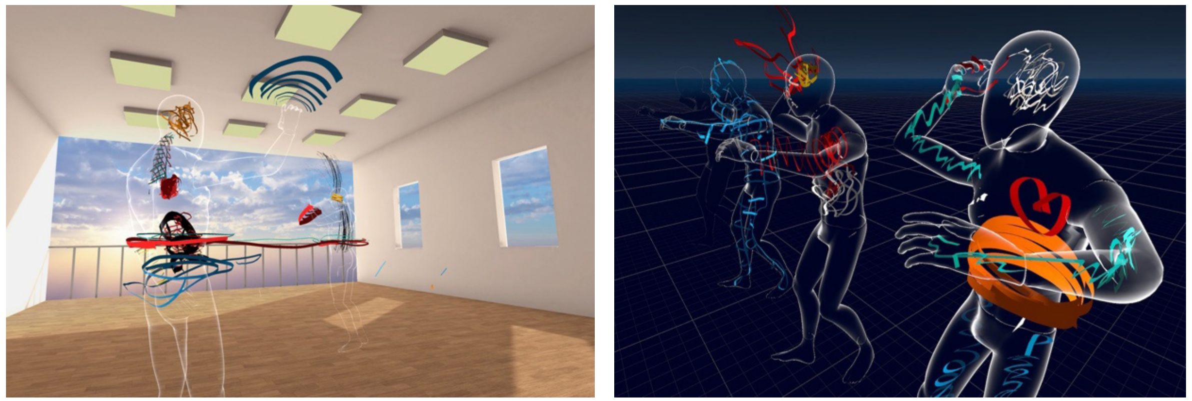

Fig. 17 Results of a drawing session created in VR. *EmbodiMap (2020).*

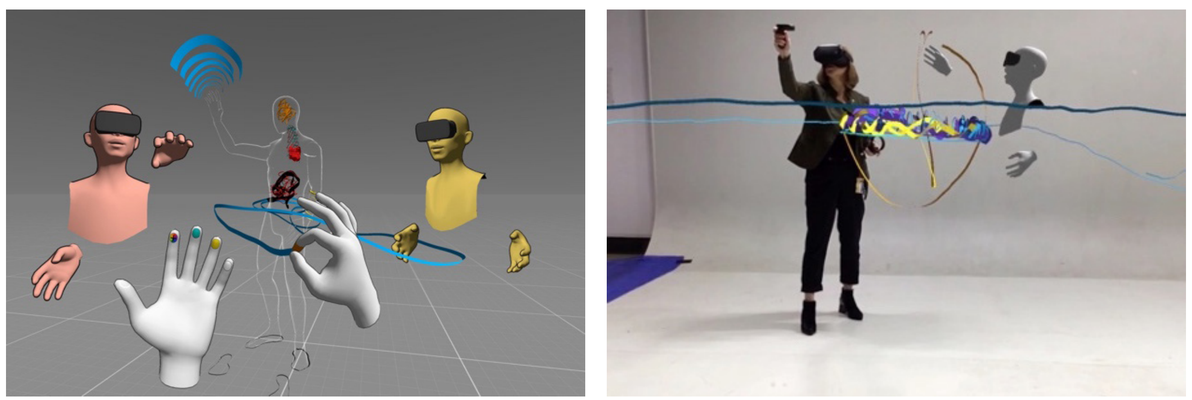

Fig. 18 Co-presence and hand interaction (left). Observer view in Augmented Reality (right). *EmbodiMap (2020).*

(Fig. 18 left). The viewer can choose a self-directed session or therapeutically guided mindful meditation. Participants can select from a range of 3D avatar bodies and shape the posture by mirroring their own pose. Various stimulating or calming environments are available to choose from and the participant can interact and draw directly with their hands. A session can be streamed wirelessly to an external monitor or as an Augmented Reality app to a tablet computer or mobile phone (Fig.18 right). *EmbodiMap* operates on the untethered portable Oculus Quest VR headset.

Unlike the previous projects, *EmbodiMap* is not based on a narrative nor does it have a linear progression (besides a guided meditation mode). As a 6DOF experience, it encourages the

viewer to explore and create. Additionally, it provides a shared experience with multiple users in the same virtual space. It is a communication tool, which is potentially valuable in the current situation of social distancing. The virtual sets are mostly abstract spaces, to help focus participants on internal sensations.

The Visit

2019, 13-15 min, language: English.

Narrative: *non-linear;* Point-of-view: *observer;* Agency: *explore (6DOF);* Cast: *virtual human;* Production: *3D geometry;* Stereoscopic: *stereoscopic;* Field-of-view: *360º spherical;* VR Platform: *6DOF;* Audio: *localised/ mono.*

The Visit emerged from research conducted by artists and psychologists at the UNSW felt Experience and Empathy Lab (feel), working with a number of women with dementia. Visitors are invited to sit with Viv, a life-sized, photorealistic animated character whose dialogue is created largely from verbatim interviews, drawing us into a world of perceptual uncertainty, while at the same time confounding stereotypes and confronting fears about dementia. The characterisation has scientific validity but also the qualities of a rich, emotion-driven film narrative. The point of the work is to draw the viewer into the emotional and perceptual world of the character, and ultimately, to break down stigma that exists around mental health.

The Visit is constructed as a real-time computer-generated interactive film, utilising virtual production methods. The character and set are represented as digital 3D models, created by employing body and spatial 3D scanning. The characters' movements are driven by motion-capture in combination with inverse kinematics. A facial rig allows her to express a wide range of emotions and to speak in sync with the voice recording. While the action and narration are scripted, she also has some autonomy, responding to the viewer's presence by making eye-contact and various facial expressions. Each experience for a viewer is unique; repeated viewing will present a different character performance and narrative strand. The environment the character inhabits is based on a

Fig. 19 Digital human character 'Viv' in her kitchen (left). The virtual set, a 3D model of a domestic kitchen, generated with structure from motion (right). *The Visit* (2019).

real location, a former home of one of the study participants. The set is comprised of a kitchen, a living room, furniture and various props, digitised utilising structure from motion and LIDAR 3D scanning.

The Visit is created exclusively with digital assets and rendered in real-time. As a consequence, a viewer is able to explore the space freely. However, the narrative invites the viewer to sit with the character at a kitchen table. At one point in the script, the character leaves to make a phone call in an adjoining room, and from observation, this is a cue for the viewer to get up and follow her.

References

Adelson E. H., & Bergen J. R. (1991). The Plenoptic Function and the Elements of Early Vision. *Computational Models of Visual Processing*, 91(1), 3-20.

Bailenson, J.N. (2018). Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do. New York: W.W. Norton.

Brozzoli C, Makin TR, Cardinali L, et al. (2012). Peripersonal Space: A Multisensory Interface for Body–Object Interactions. In: Murray MM, Wallace MT. *The Neural Bases of Multisensory Processes*. CRC Press/Taylor & Francis. Chapter 23.

Bufacchi, R. & Iannetti, G. (2018). An Action Field Theory of Peripersonal Space. *Trends in Cognitive Sciences*. 22(12). 10.1016/j.tics.2018.09.004.

Chun, W., (2016) VR: Weirder than Reality. Kenote Art&Code, Weird Reality, CMU. https://vimeo.com/214713185

Cutting J. E., & Vishton P. (1995). Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. Epstein W. , Rogers S. *Perception of Space and Motion*. Academic Press.

Davis, E. (1998). Techgnosis: Myth, magic + mysticism in the age of information. Three Rivers Press, New York.

De Jager A., Tewson A. L., Ludlow K., Boydell K. (2016). Embodied Ways of Storying the Self: A Systematic Review of Body-Mapping. *Forum Qualitative Sozialforschung*, 17(2).

Diemer J., Alpers G., Peperkorn H., Shiban Y., Mühlberger A. (2015) The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in Psychology*, 6(26).

Friedberg, A. (2006). The Virtual Window: From Alberti to Microsoft. MIT Press.

Gibson, J. J. (1997). The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.

Grau, O. (2003). Virtual art: from illusion to immersion. MIT Press.

Iñárritu, A. G. (2017). Carne Y Arena (Virtually Present, Physically Invisible). Press Release Legendary Entertainment. http://www.fondazioneprada.org/wp-content/uploads/1-Carne-y-Arena_Fondazione-Prada_press-release.pdf.

Jones, N. (2018). The Expansive and Proximate Scales of Immersive Media. *International Journal on Stereo & Immersive Media*, Vol. 2 No. 2.

Lee, K.M. (2004) Presence, explicated. *Communication Theory*, 14, 27-50.

Mine M., Brooks F. P. (1997). Moving objects in space: exploiting proprioception in virtual-environment interaction. ACM SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques.

Rizzolatti G., Fadiga L., Fogassi L., Gallese V. (1997). The space around us. *Science*, 277(5323), 190-1.

Sherman, W. R and Craig, A. B. Understanding Virtual Reality: Interface, Application, and Design. Morgan Kaufmann, Second edition, 2018.

Slater M., Wilbur S. (1997). A framework for immersive virtual environments five: Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6(6).

Project References

Being Debra. Year: 2019, platform: VR 3DOF Oculus Go, running Time: 12:05, language: English. Developed by Debra Keenahan and Volker Kuchelmeister for the Big Anxiety Festival, Sydney 2019, produced by UNSW felt Experience and Empath Lab (feel). https://kuchelmeister.net/portfolio/being-debra/

EmbodiMap. Year: 2020, platform: VR 6DOF Oculus Quest, duration: variable. Created by UNSW felt Experience and Empathy Lab (fEEL). https://kuchelmeister.net/portfolio/embodimap/

The Quiet Discomfort. Proof of concept. Year: 2018, platform: VR 6DOF Oculus Quest, 12 min, language: English. Writer/Director: Rosie Dennis, Urban Theatre Projects and Volker Kuchelmeister. https://kuchelmeister.net/portfolio/the-quite-discomfort-vr/

The Visit. Year: 2019, platform: VR 6DOF Oculus Quest, 13-15 min, language: English. Created by UNSW felt Experience and Empathy Lab (fEEL). https://kuchelmeister.net/portfolio/the-visit/

The Virtual (Reality) Museum of Immersive Experiences. Year: 2018, platform: VR 6DOF HTC Vive, running time: variable. Volker Kuchelmeister. https://kuchelmeister.net/portfolio/virtual-immersion/

Through the Eyes of the Mawa. Proof of concept. Year: 2020, platform: VR 6DOF Oculus Quest, running time: 6min, language: English. Jeffrey Aniba Waia, Andrew Belletty, Volker Kuchelmeister. Preview: https://kuchelmeister.net/portfolio/mawa/

waumananyi: the song on the wind. Year: 2019, platform: VR 6DOF Oculus Quest, running Time: narrative 12:46 - coda 10:10, language: Pitjantjatjara/English. Shot on Anangu homelands, Northern Territory Australia. Created and produced by Uti Kulintjaku, NYP Woman's Council and UNSW felt Experience and Empathy Lab (feel) for the Big Anxiety Festival, Sydney 2019. https://kuchelmeister.net/portfolio/waumananyi/