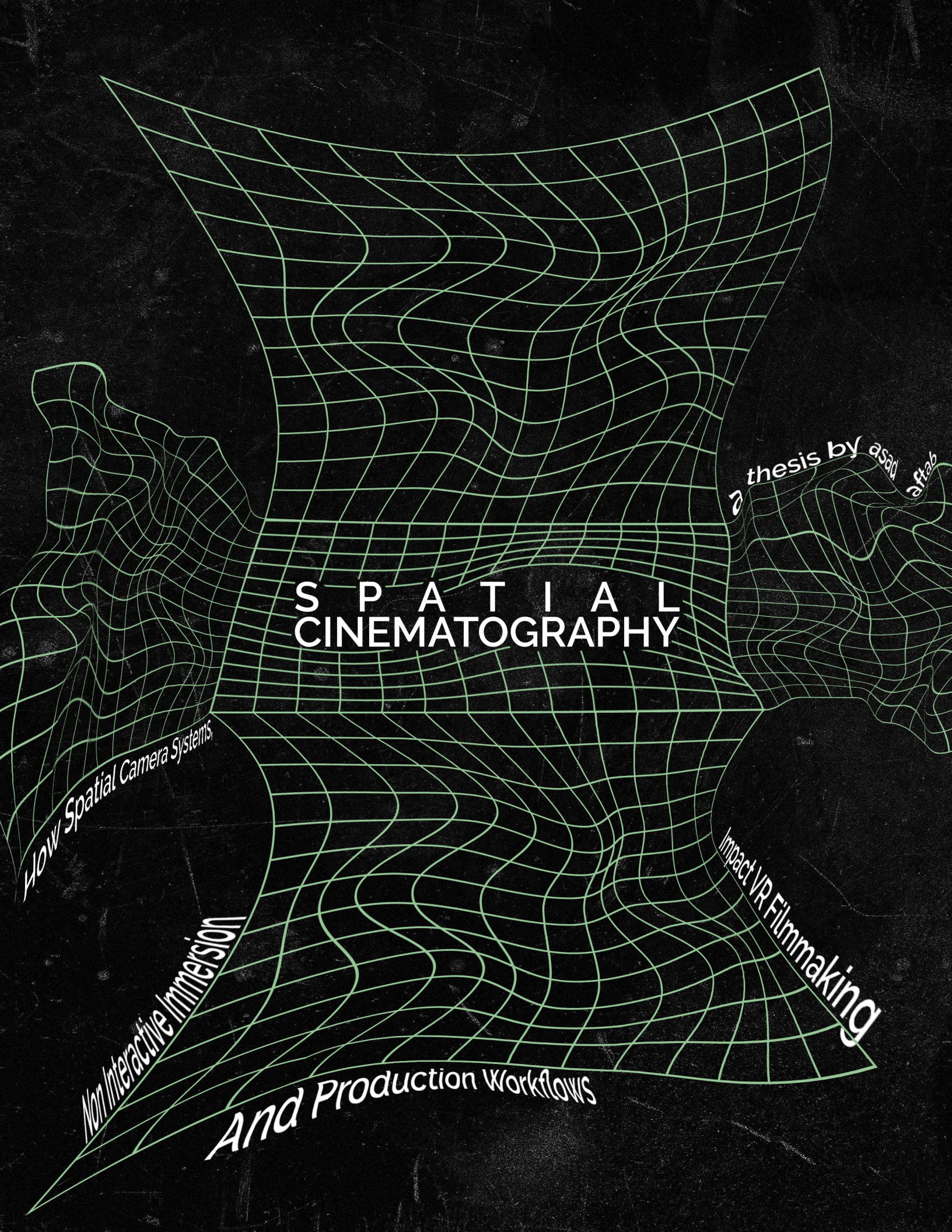




SPATIAL CINEMATOGRAPHY

How Spatial Camera Systems

Non Interactive Immersion

And Production Workflows

Impact VR Filmmaking

a thesis by asad aftab

Spatial Cinematography: How Spatial Camera Systems, Non-Interactive Immersion
and Production Workflows Impact VR Filmmaking

By

Asad Aftab

Bachelor of Industrial Design, School of Art, Design and Architecture, 2023

Supervisor: Dr. Garnet Hertz

A CRITICAL AND PROCESS DOCUMENTATION THESIS PAPER SUBMITTED IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF DESIGN

EMILY CARR UNIVERSITY OF ART + DESIGN

2025



© Asad Aftab, 2025

Table of Contents

Acknowledgments	3
Abstract	4
Keywords	4
Glossary.....	5
Introduction	7
Document Overview	7
Scope and Limitations	7
Chapter 1: Spatial Cinematography.....	9
Chapter 2: The Evolution of VR Cinema.....	11
2.1 360-Degree Video and its Evolution.....	11
2.2 The Transition to 180-Degree Video	14
2.3 Stereoscopic 3D and Cine-VR.....	16
2.4 The "Spatial Video" Revolution.....	17
Chapter 3: Pre-Production	18
3.1 Immersive vs. Interactive Content.....	18
3.2 Structuring Non-Linear Narratives	20
3.3 Pre-Visualization	22
3.4 Guiding Audience Attention.....	23
Chapter 4: Production	31
4.1 Field of View and Perspective	31
4.2 Choosing a Camera System	36
Chapter 5: Post-Production	41
5.1 Image Stabilization and Upscaling	41
5.2 Editing and Compositing	46
Research Output: The Spatial Cinematography Framework Poster	49
Conclusion and Future Directions	53
Bibliography:	55

Acknowledgments

This thesis is dedicated to people who have been instrumental in my life for the past two years.

To my friends JoJo, Jo, Qianxuan, and Rebecca, whose constant support allowed me to go beyond my own perceived capabilities and made me realize there is a life beyond work.

To Alan, Garnet, Peter and Sean, who provided me with an environment that allowed me to grow without fear of failure.

To Aisha, Hamza, Khansa, Maira, and Rija for always believing in me.

And lastly, while my family does not necessarily understand what I do, they have never opposed it.

This research was primarily conducted on the unceded xʷməθkʷəy̓əm (Musqueam), Skwxwú7mesh Úxwumixw (Squamish), and səłilwətaʔ (Tsleil-Waututh) territories, more commonly known as so-called "Metro Vancouver."

Abstract

Immersive filmmaking in virtual reality is a rapidly evolving method of storytelling that combines traditional cinematic approaches with immersive and interactive technologies. This thesis researches the creative and technological challenges of producing VR films, particularly Spatial Camera Systems, Non-Interactive Immersion, and production Workflow. By dissecting these features through a role-specific lens, this thesis guides VR filmmakers, writers, designers, and technical teams, inspiring them to explore VR's innovative potential. All of this forms the basis for the term "Spatial Cinematography."

This study emphasizes VR's creative capabilities, such as creating multi-branching storylines, assessing the level of interactivity, 3DOF vs. 6DOF immersive cinematic experiences, and using new media techniques, such as spatial sound, to invoke emotional responses from viewers. It also discusses the significance of directing audience attention in an immersive space while maintaining viewer autonomy and agency, a fundamental difficulty specific to immersive media.

On the technical side, the study examines the selection of VR-specific camera systems, image stabilization, and upscaling techniques ensuring smooth performance across various HMD platforms. It also offers best practices for production planning and efficient production workflows, drawing on practical insights from research and case studies.

Actionable solutions for expanding the field of VR filmmaking can be accomplished by evaluating current limits – such as hardware constraints and accessibility issues – and current trends like AI-driven workflows for immersive experiences.

This thesis aims to inspire filmmakers to embrace VR's potential and foster innovation and storytelling techniques that profoundly redefine audience engagement in this new, immersive narrative form.

Keywords

Keywords: *Spatial cinematography, Virtual reality, Filmmaking, Immersive cinema, AI, Immersive Video, Spatial Video, Cinematic Virtual Reality.*

Glossary

Head-mounted display (HMD): A head-mounted display (HMD) is a device you wear on your head, featuring a small optical display in front of one eye for monocular HMDs or both eyes for binocular HMDs. Virtual reality (VR) headsets are a specific kind of HMD that tracks the user's 3D position and orientation, creating an immersive virtual environment.

Stereoscopic 3D (S3D): This method crafts a depth illusion by showing a pair of images to the viewer's eyes, leading the brain to interpret them as one cohesive 3D image. The images are composed of a left image and a right image to get the two "views".

Monoscopic: It refers to a single image or video that is viewed with one eye. In this specific context, it refers to a type of virtual reality (VR) that uses a single image for both eyes.

Immersive Video: This type of video content is designed to make viewers feel like they are inside the video. The idea is to give viewers a lifelike perspective, usually in a 180-degree field of view.

Spatial Video: It is a S3D video format that allows users to move around in the video and interact with objects in it.

Pre-Visualization: Commonly referred to as Previs, this term denotes the visualization of scenes or sequences from a film before shooting begins. Pre-visualization typically includes techniques like storyboarding, which utilize either hand-drawn or digitally created sketches / 3D scenes to outline or conceptualize film scenes.

Pixel Depth: Also referred to as color depth or bit depth, it indicates the number of bits allocated for a pixel's color representation. This measurement dictates the variety of colors that can be shown on a screen. The higher the pixel depth, the higher the color quality.

3DOF: Three Degrees of Freedom (3DoF) is a virtual reality concept that defines user interaction within a virtual environment. In 3DoF, users remain stationary, allowing them to look left, right, up, and down and pivot left and right, but they are unable to navigate through the virtual space.

6DOF: In Six Degrees of Freedom (6DoF), users maintain the three types of movement enabled by 3DoF while gaining additional movement options. They can move forward and backward, up and down, and left and right. Users can also observe and interact with objects in the environment as if they were real.

NeRF: A neural radiance field (NeRF) utilizes deep learning to generate a three-dimensional representation of a scene from two-dimensional images. Put simply, it is a machine learning technique that employs AI to produce 3D models and meshes. It is a more recent advancement in 3D scanning and photogrammetry.

Lens masks: A lens mask is a video mask created around VR footage to smoothen out the edges and make it blend better into the virtual environment.

Horizon correction: This is the process of 'straightening' the footage to be level with the horizon line due to discrepancies while recording.

Stereoscopic stitching: This technique combines multiple images/videos to create a stereoscopic panorama. It is utilized in virtual reality (VR) and 3D video.

Stereo depth budget: A stereoscopic depth budget is the range of depth that a viewer can comfortably perceive in a 3D image. It is the amount of space on the z-axis that a viewer can comfortably see. The geometry of the stereoscopic pipeline from the scene to the viewer's eyes determines the depth budget.

Image stabilization: This technique introduces a compensating movement to keep the image static on the camera's sensor. This movement may be physical, i.e., In-body image stabilization (IBIS), or calculated via sensors, i.e., electronic image stabilization (EIS). IS crops the footage to compensate for the movement it is trying to remove.

Gyroscopic stabilization: It is a type of EIS that uses gyroscopic data captured either by the camera or lens to stabilize the footage at hand.

Upscaling: It involves enhancing the resolution of an image or video, which can be accomplished through an algorithm or may be powered by AI.

NLE: Non-linear editing (NLE) is a process that allows editors to modify a video or audio project without following a linear timeline. This means you can work on any clip in any sequence, regardless of whether it belongs at the project's beginning, middle, or end.

Full SBS: Full Side-by-Side (SBS) means the left and right views of a 3D video are transmitted at full resolution. This results in better quality but bigger file sizes. A 1080p full-SBS 3D file would have the following resolution - (1920+1920) 3840 x 1080P

Half SBS: Half Side-by-Side (SBS) means the left and right views of a 3D video are subsampled at half resolution, and you get a backward-compatible full frame. A 1080p half-SBS 3D file would have the following resolution - (960+960) 1920 x 1080P

Screen door effect: The screen door effect, or pixelation, is a visual artifact in which the spaces between pixels on a display become apparent. This effect is reminiscent of a mesh or flyscreen and is particularly evident in low-resolution displays and VR headsets.

SOOC: Straight out of Camera – This means that whatever image/video file we are looking at has had no post-processing or corrections applied to it and is being viewed exactly as the camera shot it.

Introduction

Virtual Reality filmmaking, also referred to as Cinematic VR (Mateer, 2017) or Cine-VR for short, has evolved significantly, transitioning from early 360-degree video experiments to today's new "spatial video" revolution in digital content and hardware. This thesis examines how spatial camera systems, non-interactive immersion, and production workflows shape Cinematic VR content. By understanding these three base variables, we can develop a framework for telling a compelling story in an immersive setting.

Document Overview

This document is split into five main chapters, starting with Spatial Cinematography in Chapter 1, which provides an overview and definition of Spatial Cinematography, what Cine-VR is as a medium, and the role of a Spatial Cinematographer. Next, Chapter 2 explores the Evolution of VR Cinema, explaining a historical market overview of the medium. Chapter 3 continues with pre-production, highlighting tools like pre-visualization methods, spatial audio cues, and visual design techniques to strategically guide viewer attention. Production is the theme of Chapter 4, which explains field of view, guiding viewer attention, choosing a spatial camera system and balancing the technical and creative aspects of a Cine-VR project. The final Chapter 5 finishes with an overview of post-production techniques. After this, the document shows how these principles play out in practice through an easy-to-consume and distribute poster. Moreover, it looks at the prospects of the field.

Scope and Limitations

This project intentionally limits its scope to focus on the technical aspects of VR production for spatial cinematography. It does not address the types of content that can be created within the VR format; rather, it offers practical production workflows to assist creators with the medium's formal challenges. It strives to be a practical technical manual. Broader ethical, social, and environmental concerns surrounding VR technologies are also outside the scope of this research. While important issues exist—including the biases embedded in media, as McLuhan described in his assertion that "the medium is the message" (McLuhan, 1964), as well as corporate critiques of companies such as Meta (with Mark Zuckerberg)—they are not explored here. Environmental impacts, including resource use, carbon emissions, and e-waste associated with VR systems, are similarly acknowledged but not analyzed in detail. These exclusions are not meant to diminish the significance of these concerns, but to

maintain a clear focus on addressing the lack of accessible technical knowledge in VR production. This document is therefore structured intentionally as a practical technical manual.

Chapter 1: Spatial Cinematography

I refer to this approach towards VR Filmmaking as "Spatial Cinematography," and it extends beyond technological considerations, influencing the very nature of storytelling and audience engagement. Traditional cinematography techniques must be adapted and modified for immersive experiences (Zhang & Weber, 2023), requiring us to rethink narrative construction, cinematographic framing, and viewer interaction. This research builds on existing ideas in VR Film while proposing new methods for maximizing engagement through spatial cinematography techniques.

An important distinction is that Cine-VR is a medium like Film and Digital Cinema and can be defined as being similar to 360-degree video, but its key difference from 360° is using cinematic production techniques such as lighting design, sound design, scenic design, and blocking techniques (the latter two in the case of dramatic work) (Williams et al., 2021).

Within this new medium of Cine-VR comes this newly proposed field called Spatial Cinematography and this also gives rise to a new role in the production of the Spatial Cinematographer, which is a combination of a traditional cinematographer/Director of Photography, a Stereographer and an XR designer/artist.

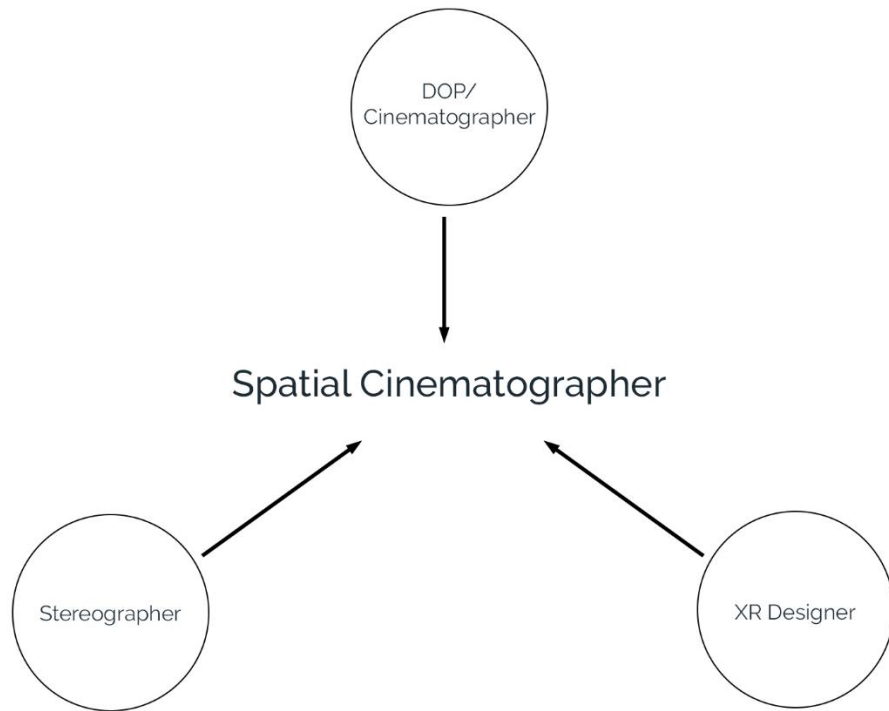


Figure 1: *Multi-sided role of the Spatial Cinematographer.*

Creating this new multi-faceted role is important as VR cinema is cementing itself as an increasingly rapidly expanding medium for immersive storytelling. This requires streamlined roles, responsibilities, and workflows to reduce the barrier of entry for new creatives who wish to use this medium to tell their stories more effectively.

Chapter 2: The Evolution of VR Cinema

2.1 360-Degree Video and its Evolution

The birth of 360-degree video introduced audiences to an entirely new way of viewing content. However, the format presented challenges, such as limited storytelling control and technical difficulties in camera placement and stitching.

In the early 2000s, the novelty of 360-degree filmmaking drew early adopters, but its constraints soon became evident. Initially, multiple cameras had to be used in custom-made enclosures to film a 360-degree video. The most common example of this was the usage of GoPro action cameras in an enclosure like the one in Figure 2 below.

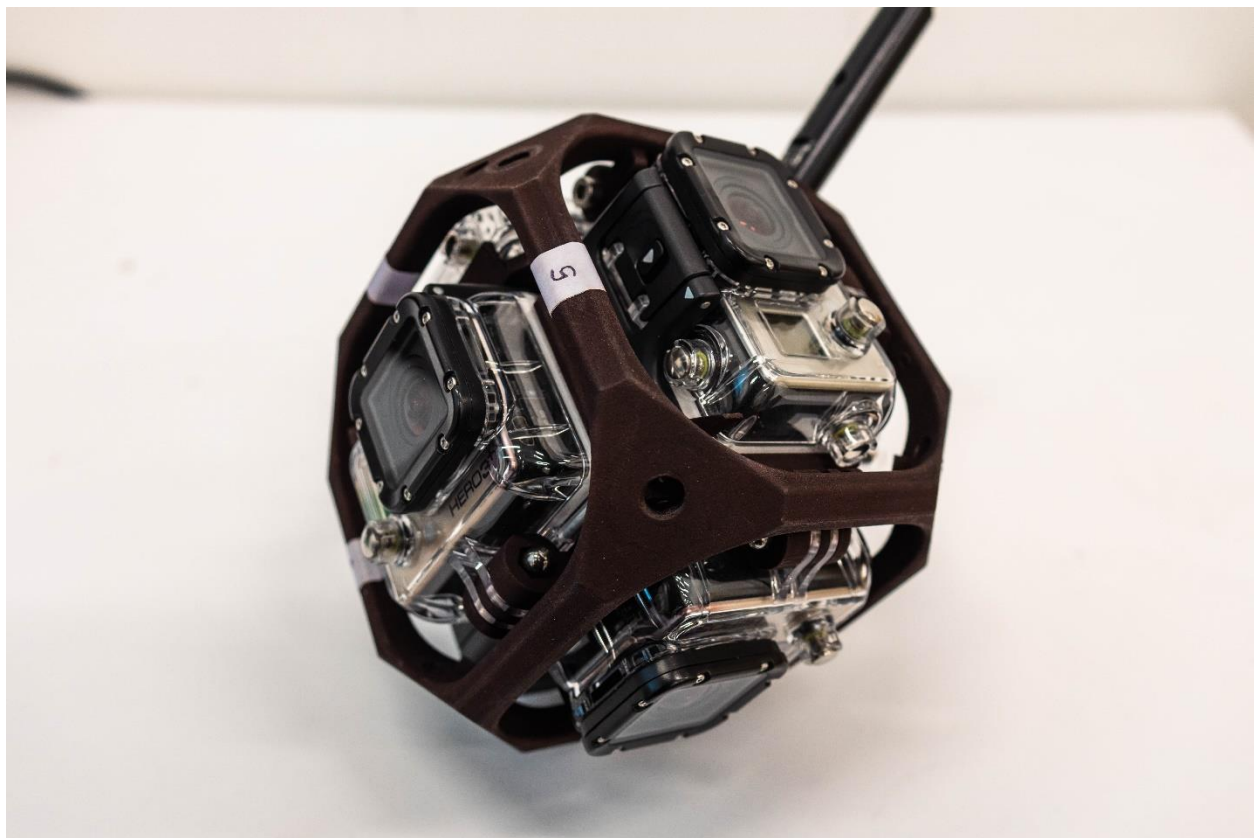


Figure 2: Custom GoPro enclosure made to hold 6 GoPro Hero 3+ to film 360 footage from 2012.

The most significant disadvantage of this GoPro array approach was the lack of image stabilization and manual stitching together in posts. The results ranged from decent to downright abysmal, but it was a start. As the years went by, and 360 cameras (and VR headsets) got better, we started to see more polished solutions from companies like

Insta360 and Ricoh in the form of Dual-fisheye lens cameras in 2013, which allowed filming in 360° to be much simpler. The introduction of auto-stitching software helped to speed up editing and make 360° videos more accessible to the average consumer.



Figure 3: *Insta 360 ONE (2017).*¹

At the high end, we also saw companies like Kandao and Insta360 introduce "pro-level" cameras in 2017 and 2021. These cameras can shoot at high resolutions of 8K and beyond and have many features to streamline 360-degree filming and live streaming.

¹ Read more about the Insta360 ONE: <https://www.cnet.com/reviews/insta360-one-preview/>



Figure 4: Left Insta 360 Pro² from 2017, right: Kandao Obsidian Pro³ from 2021.

However, 360° videos since the early 2000s have some major flaws that keep them from reaching their true potential, which primarily boils down to the fact that the camera captures everything around them. Lighting scenes using grip and lighting gear becomes extremely difficult and, in many cases, impossible (Chulhyun, 2016). Coupled with the difficulties of framing and the increasingly difficult challenge of keeping the audience engaged in whatever narrative was being strung, it made life difficult in the early days of VR cinema. It resulted in visually stunning pieces but fell flat in other areas. Eventually, 360° videos were relegated to a small niche as they began to be viewed as something that required extreme effort to produce and often resulted in very limited audience engagement.

This prompted more interactive experiences to be realized using tools like Unity and Unreal Engine, and they like to take centre stage in VR content. Although these developments reduced passive viewing and enhanced immersion, they failed to fulfill the promise of 'Cinematic VR.' Instead, they provided gamified experiences that required constant user engagement and physical interaction, complicated by the discomfort of bulky headsets.

² Read more about the Insta 360 Pro <https://www.insta360.com/product/insta360-pro>

³ Read more about the Kandao Obsidian Pro <https://www.kandaovr.com/Obsidian-Pro>

These challenges with the 360°-video format between 2012 and 2020 prompted VR filmmakers to seek more engaging formats. As a result, several content creators began experimenting with 180-degree videos around 2019.

2.2 The Transition to 180-Degree Video

As interest in 360° started to wane, the industry started to look for other formats to create Cinematic VR content. In around 2018, Google released the VR 180° format (Ackermann, 2018) as a more usable alternative to 360°.

The Lenovo Mirage was the first consumer-facing VR 180° camera to hit the market in 2018 (MIXED, 2023), soon to be followed by offerings such as the Insta360 EVO and Vuze XR, which still remain popular among enthusiasts.



Figure 5: Left: Vuze XR from 2019⁴, Right: Insta 360 EVO from 2019⁵.

Unlike 360° cameras, 180° cameras quickly emerged after around 2019 as a more practical alternative for Cinematic VR despite its initial struggles to gain traction due to various technological constraints. However, improvements in hardware and software, largely thanks to the introduction of Canon's EOS VR system, have made 180-degree Video a viable option for immersive storytelling, balancing a controlled perspective with a heightened sense of depth and immersion,

⁴ Read more about the Vuze XR <https://www.videomaker.com/reviews/cameras/humaneyes-vuze-xr-hands-on-review-a-flexible-camera-for-the-cutting-edge/>

⁵ Read more about the Insta 360 Evo <https://www.insta360.com/product/insta360-evo>



Figure 6: Canon R5 paired with the RF 5.2 mm F2.6 Dual Fisheye Lens, released in 2021⁶.

⁶ Read more about the EOS VR system at https://canon.ca/en/product?name=RF5.2mm_F2.8_L_Dual_Fisheye&category=/en/products/Lenses/VR-Lenses

2.3 Stereoscopic 3D and Cine-VR

Stereoscopic 3D (S3D) adds a layer of depth perception, making environments feel more tangible and "real." S3D and Cinematic VR go almost hand in hand to the point that Monoscopic cine-VR tends to look 'flat.'

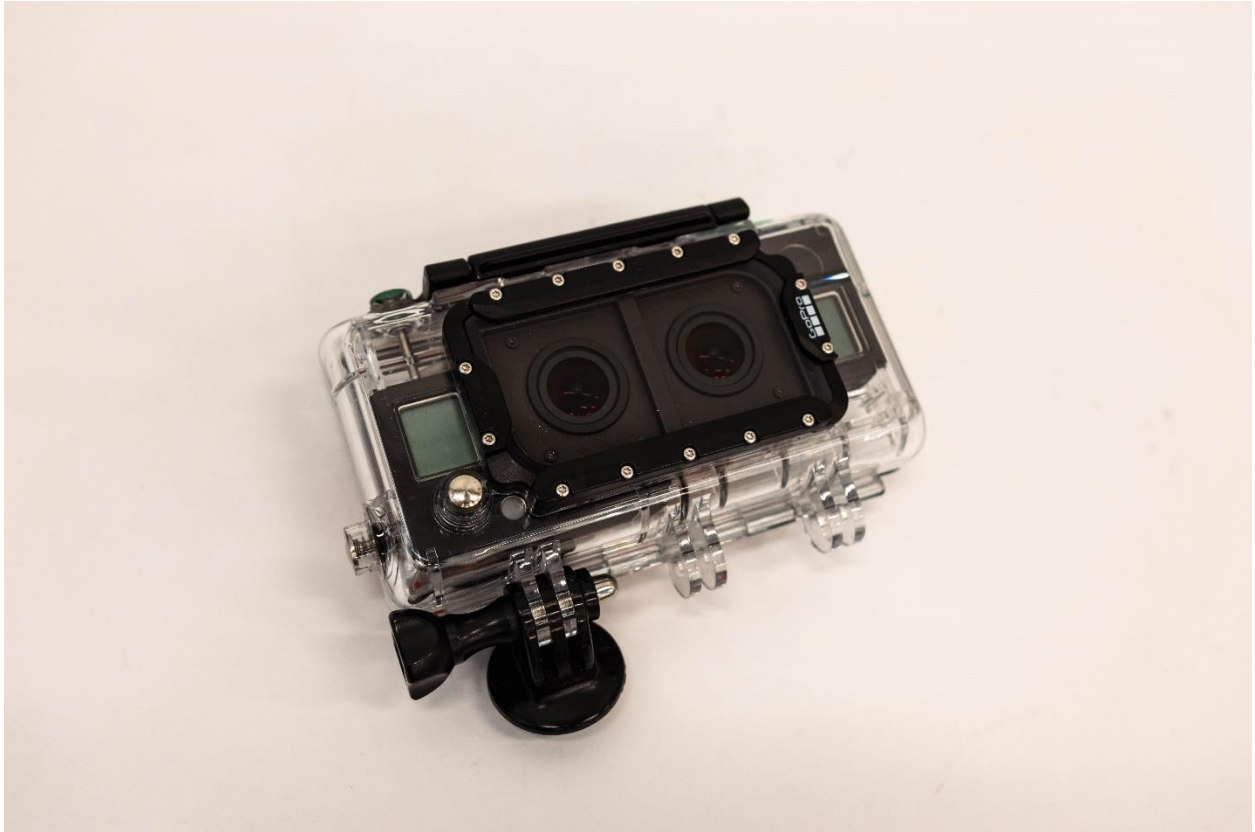


Figure 7: GoPro Dual Hero System⁷ – A popular consumer-level S3D platform released in 2015 (Asch, 2022).

Stereoscopic 3D (S3D) became a popular format in the early 2010s, championed by filmmakers like James Cameron and his groundbreaking 'Avatar.'(CITE) However, S3D is historically a very old medium, being especially popular in the cinema space since the 1950's (Arden, 2012). Unfortunately, difficulties in filming in 3d and the high cost of entry for consumers eventually prevented the format from widespread adoption. (Peddie, 2024)

However, as VR headsets improved in screen resolution and pixel depth (Lynn et al., 2020), they almost resurrected 3D as a format outside of theatre screens. You now do not need a

⁷ Read more about the dual hero system here <https://stereoscopy.blog/2022/01/16/gopro-dual-hero-system-housing-an-extraordinary-3d-camera-rig/>

3D TV or an expensive projector and 3D glasses setup. You just need a semi-decent VR headset for a more than acceptable 3D viewing experience.

However, filming for VR in S3D also presents its own unique challenges, such as maintaining interocular distance (Knorr et al., 2024) especially when using DIY filming solutions and avoiding eye strain due to improper stereo calibration (Terzić and Hansard, 2016). However, the effort is worth the hassle as it helps increase the viewer's immersion, and the layer of depth perception brings it closer to real life (Morana, 2024).

2.4 The "Spatial Video" Revolution

In December 2023, Apple introduced 'Spatial video' on the iPhone 15 Pro, which they termed "A groundbreaking new capability that helps users capture life's precious moments." (Apple Newsroom, 2023). This new format of immersive video capture was brought forward on the heels of the launch of the Apple Vision Pro. Marketing jargon aside, "Spatial video" is just Stereo 3D video, and "Apple Immersive Video" is simply a 3D VR 180° video (Swanson, 2024).

With the release of the Apple Vision Pro, we are seeing a big shift in the Cine-VR industry. One of the most significant issues plaguing standalone VR headsets is the lack of powerful hardware to run high-resolution screen panels and playback high-resolution footage. The AVP solves this by effectively being a MacBook Pro strapped to your face. With the creation of the new MV-HEVC codec for 'spatial video,' (Blackman and Harley, 2024) We are finally seeing a clear push towards high-level Cine-VR content. With Meta also supporting this new video container on their headsets, it would not be remiss to say that the stage is set for the next generation of VR filmmaking.

The push for high-fidelity experiences due to the 'screen door effect' (Cho et al., 2017) has driven technological innovation. We finally have both the hardware and software to make said experiences, but compared to traditional filmmaking, Cine-VR still lacks established frameworks and 'best practices' owing to its relatively nascent stage. This necessitates the formalization of 'spatial cinematography' into its own creative field within CVR.

Chapter 3: Pre-Production

This section looks at some of the more theoretical and creative decisions you will have to make when making the deep dive into the world of Cine VR. It unpacks key strategies for building compelling storytelling within virtual reality—an area often caught between traditional cinematic conventions and interactive experiences. This chapter highlights concrete tools like effective pre-visualization methods, spatial audio cues, and visual design techniques to strategically guide viewer attention. Ultimately, it outlines how filmmakers can leverage these principles practically, pushing Cinematic VR beyond gimmicks into meaningful, impactful narratives.

3.1 Immersive vs. Interactive Content

VR content/experiences can primarily be grouped into two categories based on the amount of movement freedom the user has:

- 3DOF: Head tracking around 3 rotational axes.
- 6DOF: Head tracking + translational motion tracking.

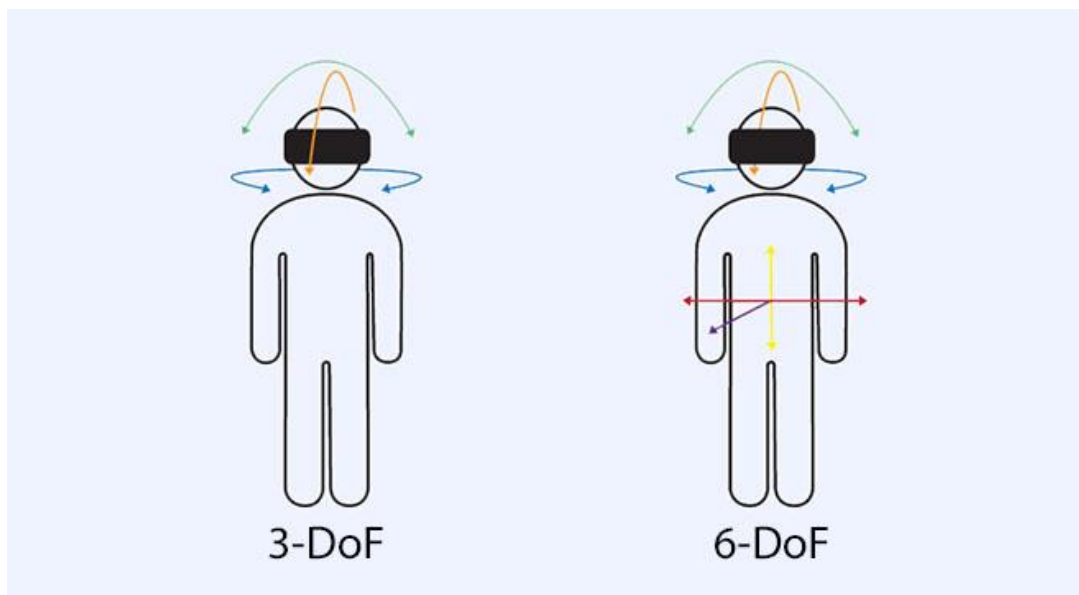


Figure 8: An illustrative example of 3DOF vs 6DOF (Barnard, 2023)

In simpler terms, in a 3DOF experience, the user is an **observer**, while in a 6DOF experience, the user is a **participant** (IEEE, 2022).

Many VR experiences we see now are primarily 6DOF, and perception has been built around the medium's requirement that the user be able to touch and interact with objects in the scene and weave their own interactive storyline. However, 6DOF experiences are closer to games than to cinema, and that is a barrier to entry for both creation and consumption, as people may not possess the necessary technical skills to make a whole 6DOF experience in Unity or Unreal Engine and similarly, if not executed correctly 6dof experiences could potentially lead to more immersion breaking scenarios than 3DOF (Chatterjee et al., 2024).

It is almost universally agreed upon in the VR community that 3DOF is a "lower level" of immersion than 6DOF (IEEE, 2022) However, in the case of Cine-VR, creating a live-action 6DOF experience is incredibly difficult/near impossible due to the sheer amount of resources required to film all the possible paths the user may take (which are practically infinite). So, for Cine-VR specifically, 3DOF is the sub-medium of choice, and 'non-interactive immersion' is required within that container.

Non-interactive immersion allows for greater directorial control while maintaining audience engagement. As we have a 3DOF-esque experience, we eliminate the possibility of the user being able to interact with the environment actively. Previous studies have found that, generally, viewer attention is focused on the centre area of the frame (L. Tong et al., 2021). This makes the shift to VR 180° even more sense as it allows the content to be framed in a more user-consumption-friendly way while maintaining a high immersion level. If screenwriters and directors from traditional films have experience with formats such as Vista Vision or IMAX, they will find writing a story for immersive to be similar to those two formats as they are primarily "centre-punched". (IMAX Corporation, 1999)

Therefore, by carefully designing non-interactive experiences, filmmakers can maintain a higher level of agency over the narrative than usual while still offering audiences the sensation of presence. In other words, this research suggests that 3DOF is a considerably superior choice over 6DOF as an immersive video narrative format. This goes against the assumption that "more is better" for degrees of freedom: for CineVR, three is better than six.

3.2 Structuring Non-Linear Narratives

Non-linear storytelling in VR challenges traditional narrative structures by introducing audience interactivity and spatial exploration. Unlike conventional filmmaking, where scenes unfold in a predetermined sequence, VR narratives must account for user agency while maintaining coherence. This balance requires innovative storytelling techniques such as environmental cues, guided focus, and adaptive pacing.

As VR blurs the lines between linear and non-linear storytelling, immersive experiences like *CARNE y ARENA* (2017) demonstrate how spatial narratives can engage viewers while preserving artistic intent. To navigate this complexity, VR filmmakers rely on advanced pre-visualization tools and spatial storyboarding to direct attention and enhance narrative immersion.

Due to its inherently non-linear nature, virtual reality (VR) presents unique challenges and opportunities for storytelling. Unlike traditional filmmaking, where the narrative unfolds in a fixed sequence, VR and immersive media allow for greater audience interactivity, creating a complex storytelling environment where guiding viewer attention becomes critical. (Schleser et al., 2024)

While VR is often categorized as a non-linear medium due to its interactive capabilities, immersive and spatial videos occupy a middle ground between linear and non-linear storytelling. In these formats, directors and screenwriters still control the overarching narrative. However, due to the expansive field of view in VR—particularly in 360-degree video—the audience has the freedom to explore the environment, which can lead to divided attention (Choi and Nam, 2022). In 360-degree experiences, the viewer's attention is divided between environmental exploration and the primary narrative elements. This creates a significant challenge in ensuring that key narrative moments are not overlooked.

A stellar example of non-linear narratives in VR is the Oscar special award-winning VR installation (AMPAS, 2017) "**CARNE y ARENA (Virtually present, Physically invisible)**" by director **Alejandro G. Iñárritu**.

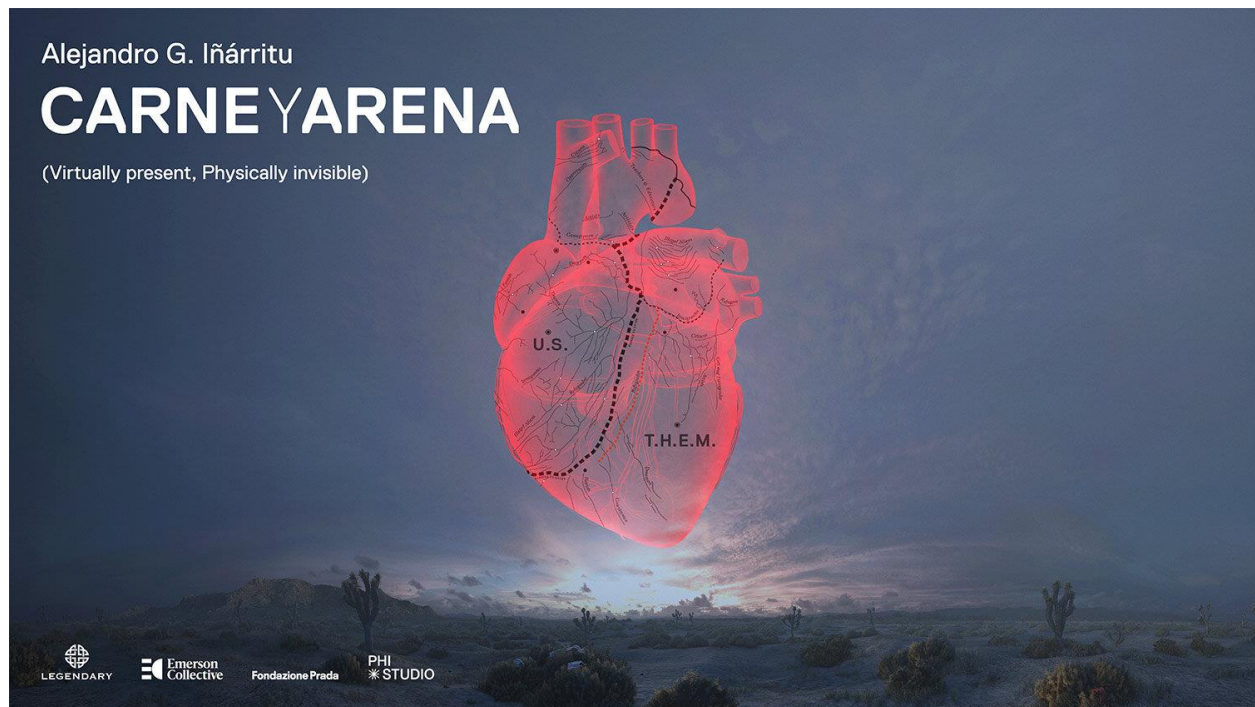


Figure 9: Title poster for *CARNE y ARENA* (PHI, 2017)

CARNE y ARENA explores the human condition of immigrants and refugees. This is a twenty-minute individual experience focused on a virtual reality segment shared by three guests in distinct rooms. Each guest is able to experience the story at the same time from their own perspective.

To effectively structure such an experience, the spatial cinematographer must employ strategic pre-production planning and visualization techniques to mitigate the risk of audience distraction and ensure that critical narrative elements are perceived. In traditional filmmaking, storyboarding is used to plan shots and sequences. In VR storytelling, storyboarding through the use of pre-vis techniques like using Unreal Engine 5 becomes even more essential, as it must account for the audience's ability to look in multiple directions (Gipson et al., 2018). Therefore, storyboarding in VR aids production logistics and serves as a tool for guiding user focus.

VR storytelling demands a meticulous approach to narrative structuring, balancing the freedom of audience exploration with the need to convey a cohesive story. Through careful pre-production planning, effective use of audiovisual cues (Vosmeer and Schouten, 2017), and strategic set design, creators can successfully navigate the challenges of non-linear storytelling in immersive media. By leveraging these techniques, VR filmmakers can create

engaging and meaningful narratives that resonate with their audiences while maintaining coherence within an expansive, interactive space.

3.3 Pre-Visualization

Previsualization, or pre-vis for short, plays an increasingly important role in modern filmmaking (Ardal et al., 2019). However, it remains less prevalent in smaller productions than traditional storyboarding techniques. In CVR, pre-vis is particularly critical because it enables filmmakers to simulate the virtual environment and test audience reactions before full-scale production begins (Park, 2018).

By utilizing game engines such as Unreal Engine, filmmakers can create interactive pre-vis models that allow them to assess viewer focus (Muender et al., 2018). For instance, user testing can reveal whether audiences are paying attention to key narrative elements or becoming distracted by peripheral details in the scene. These insights inform staging, lighting, and production design adjustments to enhance storytelling clarity.

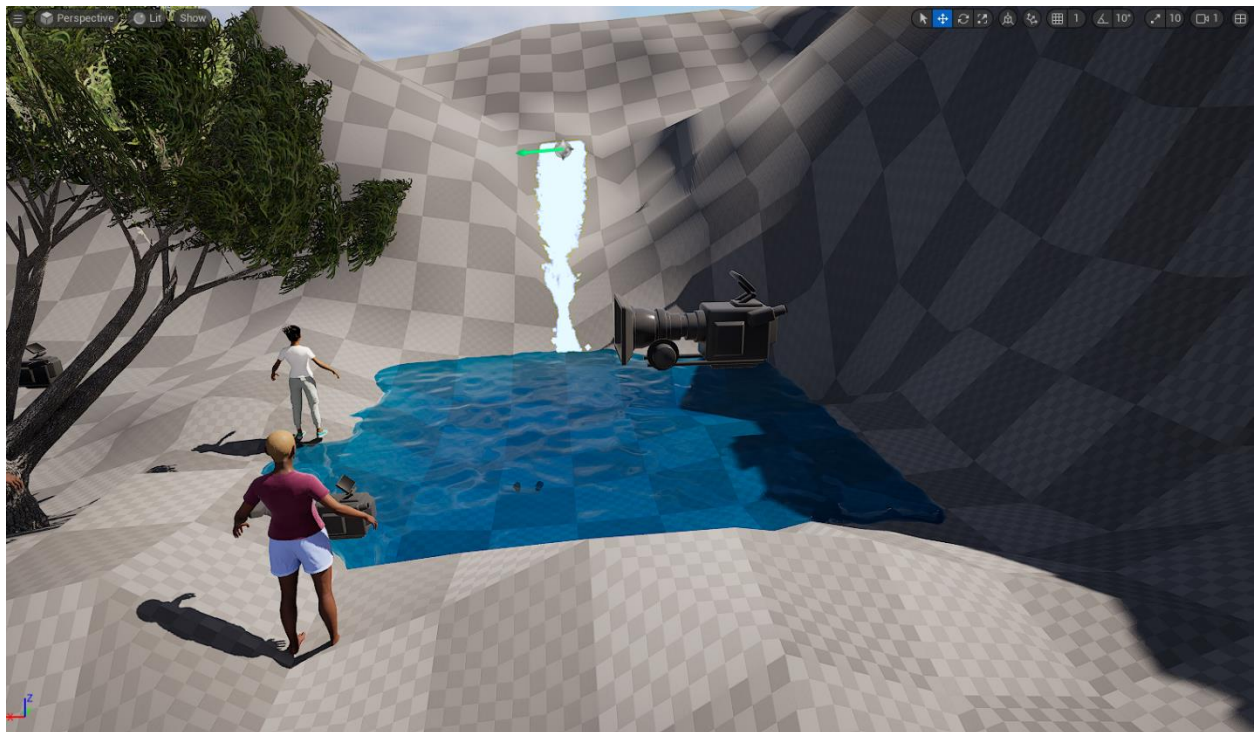


Figure 10: A pre-vis example in UE5.

Given the high production costs associated with CVR, allocating sufficient time and resources to both previsualization and user testing within the pre-production pipeline is essential. Unlike traditional films, CVR projects demand a more extensive development

process due to the medium's complexity, the integration of interactive elements, and the unique distribution challenges.

Newer techniques of 3D scanning like NeRF's combined with VR (Xu et al., 2023) have opened up many possibilities for recreating real-life locations for the purposes of Previs. Thanks to apps like Luma AI, NeRFs have become much more accessible. They allow one to simply use their phone to scan the room and either cloud process or locally process it on their own PC with very convincing and usable results.

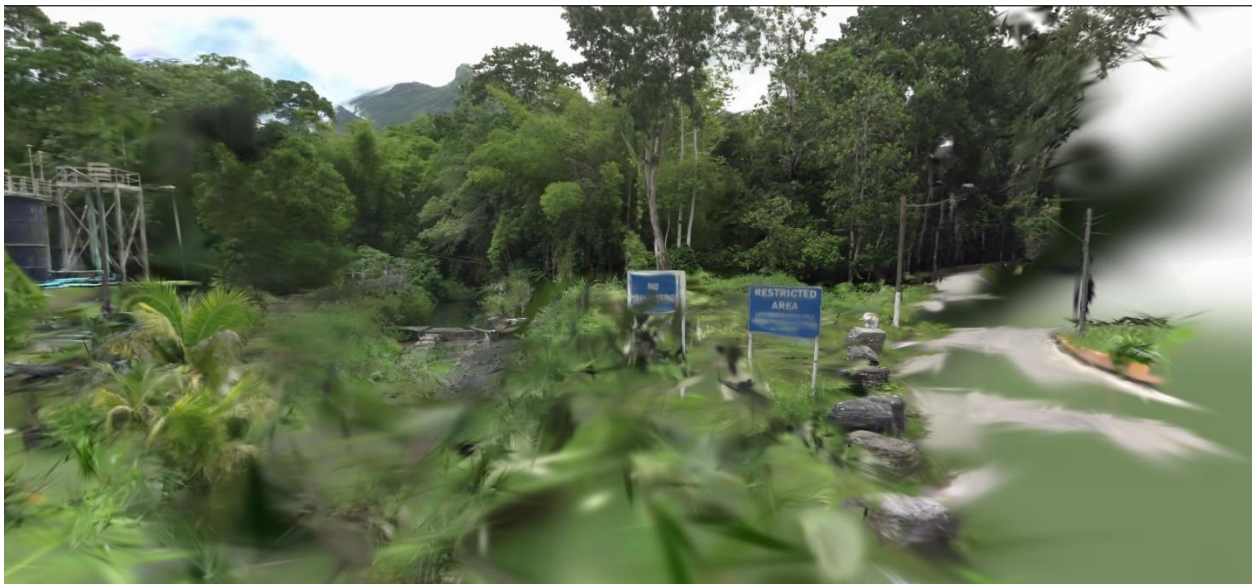


Figure 11: An example of a still screenshot from a NeRF scan; view the scan at this [link](#).

3.4 Guiding Audience Attention

In VR Cinema, audience focus is more difficult to direct than in traditional cinema. Techniques such as spatial sound cues, lighting, and environmental design are crucial in maintaining immersion without explicit direction. Ensuring that viewers naturally follow the intended flow of the story is essential to sustaining engagement and coherence (Knorr et al., 2018).

In the production of Cinematic virtual reality (CVR), several key elements contribute to the overall immersive experience, including sound design, lighting, environmental design (i.e., set design), and production design (Choi and Nam, 2022). Unlike traditional filmmaking, CVR does not provide explicit directional cues such as arrows, subtitles, or guided prompts that are often present in interactive experiences. One of the most significant challenges in non-interactive immersive experiences, as opposed to interactive ones, is the absence of non-

playable characters (NPCs) to direct the user's attention. This lack of guidance means we do not know where the viewer will look exactly (He and Liu, 2024) and it necessitates a different approach to storytelling, requiring the director, screenwriter, and production team to anticipate user behavior and design the experience accordingly.

Several techniques can be used to direct viewer attention effectively in VR (Carpio et al., 2023):

- **Audio Cues:** Sound design plays a crucial role in VR storytelling. (Truong, 2015). Directional audio can draw attention to specific scene areas, subtly guiding the viewer's gaze. As part of the **PXR conference XRtist linkup**, I developed a VRChat world titled "**Cabin Fever**" alongside a team of students from various universities. I was the main Unity World builder on this project, and as a mechanism to guide the user, various audio cues were used to help them find the required keys to progress through the experience.

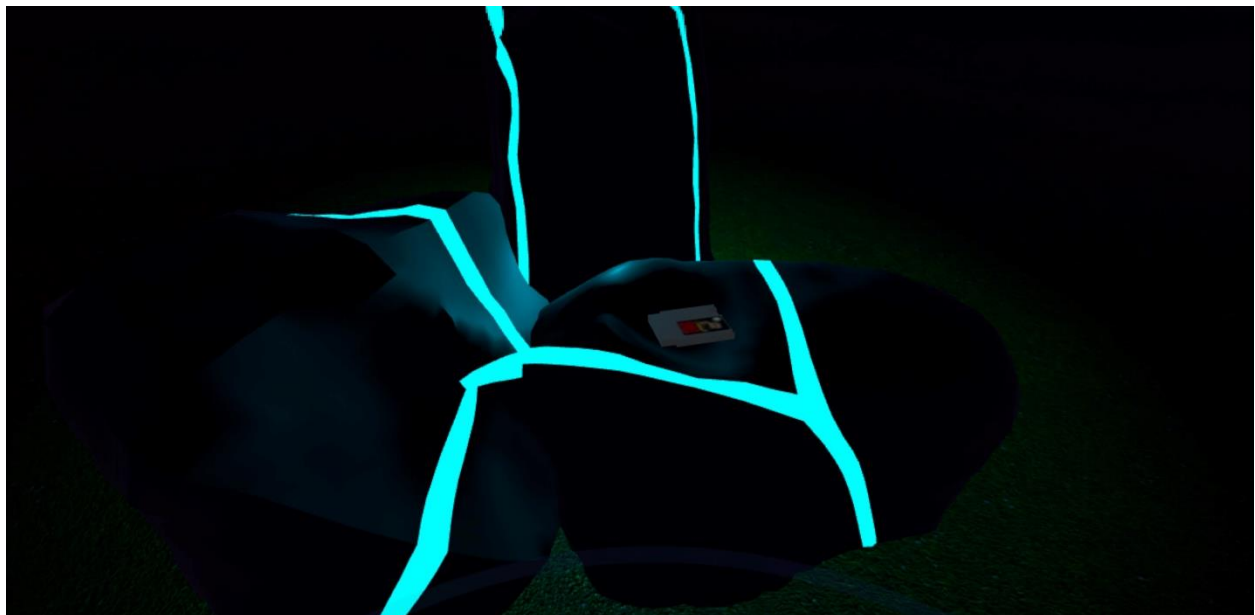


Figure 12: Screenshot from "**Cabin Fever**" – The experience can be viewed in VRChat [here](#).

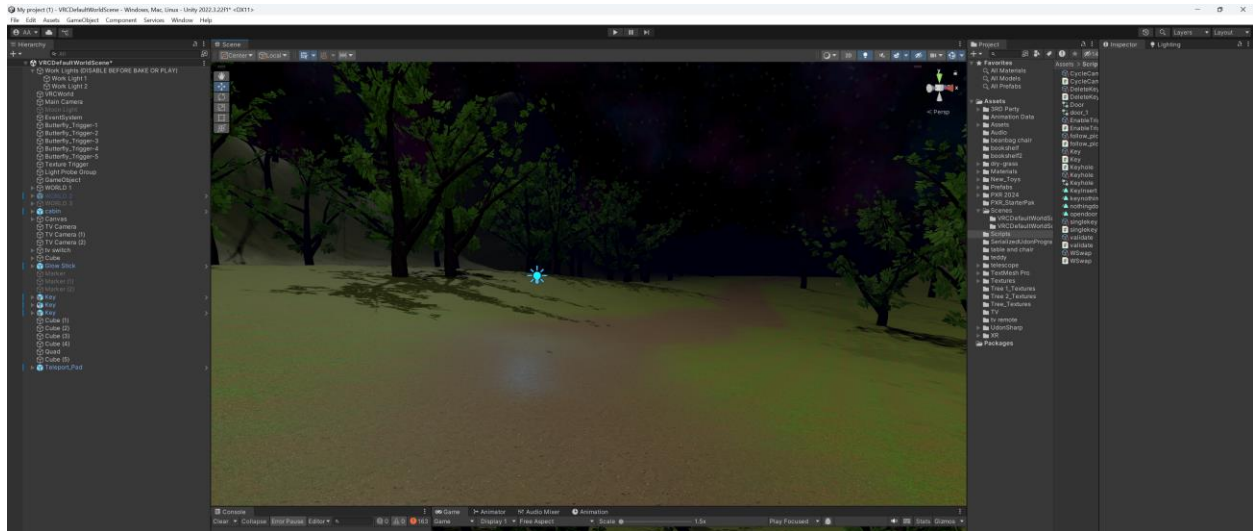


Figure 13: Unity project screen grab from "Cabin Fever".

- Gestural Cues:** Character movements and gestures can serve as natural focal points, encouraging the audience to follow the narrative action. For example. In this untitled VR 180° piece, the audience is encouraged to look around the 180°-environment based on the actions of various people in the scene and also with camera movements.



Figure 14: Screenshot from "Untitled" VR 180° piece. View it at this link [here](#).

- **Environmental Design:** Set design and production elements can be structured to lead the viewer's eye toward key storytelling elements subtly. For example, objects of interest can be highlighted using lighting contrasts or motion. In my 360 VR piece titled "**Dreamstate – Ascension**," the movement of the various spacecraft across the screen is used as a way to direct the user's eye across most of the space and to encourage them to look around the 360 environments.



Figure 15: Screenshot from "Dreamstate - Ascension". View it at this link [here](#).

- **VFX Enhancements:** Visual effects can strategically emphasize important elements within the scene, ensuring that narrative significance is reinforced even in a highly interactive space (Park, 2018). Again referencing "Cabin Fever," in this experience, the keys had a subtle glow texture, and the path to them had hints of small glowing rocks as wayfinding points. The main guiding "fairy" was modelled as a striking orb of light to instantly pull the user's attention to it as soon as they stepped into the scene.



Figure 16: Guiding fairy from "Cabin Fever".



Figure 19: Guiding Rocks from "Cabin Fever" – The experience can be viewed in VRChat [here](#).

Beyond directing attention, set and production design play a significant role in enhancing the immersive narrative experience. Given the wide frame of VR environments, every element within the scene must contribute meaningfully to the story. Whether filming in an outdoor location, like a forest, or an elaborate interior space, like a mansion, the careful set decoration can incorporate visual cues that reinforce the storyline. Additionally, post-production techniques can integrate subtle narrative elements into the environment, enriching the immersive experience and making the world feel more alive.

A crucial aspect of this whole process is user testing. In interactive experiences, interactivity is tested to ensure functionality and engagement. Similarly, in CVR, testing is essential to evaluate how viewers engage with the visual and auditory elements. Traditionally, screen tests in filmmaking occur during the later stages of production, often after a rough or final cut is available. However, in CVR, the testing process must begin much earlier, particularly during pre-vis.

Another key distinction between CVR and interactive experiences, such as video games, is the level of viewer agency (L. Tong et al., 2021). In interactive media, designers can restrict user movement or progression until specific actions are performed. In contrast, Cinematic VR often lacks such constraints, especially in exhibition settings like film festivals. While users may have limited control, such as being unable to pause the experience if controllers are removed, their gaze and attention remain unrestricted. Consequently, if crucial narrative elements are overlooked, they may be permanently missed, posing a challenge for filmmakers who seek to craft narrative-driven VR experiences (Szita and Gander, 2021). As narrative-focused CVR projects become more common, filmmakers must develop strategies to direct audience attention within the immersive environment effectively.

This challenge also underscores the advantage of VR180° over full 360-degree VR. In a VR180° experience, the viewer's field of vision is limited to 180°, reducing the likelihood of distractions and enabling the production team to focus on directing attention within a defined frame (L. Tong et al., 2021). By contrast, 360-degree VR allows viewers to look in any direction, increasing the risk of missing key story elements. As a result, VR180° provides a more controlled environment that aligns more closely with traditional cinematic storytelling techniques.

In conclusion, Cinematic VR development requires a strategic approach to production, incorporating Previs, user testing, and careful attention to audience engagement. As the field continues to evolve, these considerations will be instrumental in shaping the future of immersive storytelling.

Chapter 4: Production

Creating content for spatial cinema demands a thorough grasp of both the technical and creative aspects that are distinct to VR filmmaking. In contrast to traditional cinema, where the framing and perspective are closely managed, spatial cinematography must navigate the delicate interplay between immersion and narrative clarity. One of the crucial choices a spatial cinematographer faces is the selection of the field of view (FOV), which significantly impacts how viewers interpret and interact with the narrative.

This chapter delves into the essential elements of spatial cinema production, starting with how the field of view (FOV) and perspective influence storytelling. It continues by discussing the significance of selecting the appropriate camera system, spanning from budget-friendly options to advanced professional setups, each providing unique benefits depending on resolution, workflow, and production requirements. By grasping these key components, VR filmmakers can make informed choices that elevate both technical quality and immersive storytelling.

4.1 Field of View and Perspective

Selecting between 180-degree and 360-degree FOV is perhaps a spatial cinematographer's most crucial decision to best convey the story the director and screenwriters have created.

As previously discussed, the FOV influences storytelling and audience experience. So, the spatial cinematographer must work with the pre-production team during the pre-vis process to select the FOV that best conveys the director's vision and is practical in production.

While a wider field of view can enhance presence, environment building, and details, it may introduce complexity in directing viewer focus as a unique quality of Cinematic VR is that the viewer is generally placed into a first-person "POV" style perspective as opposed to seeing things through the lens of a camera as that tends to be more effective (Cannavò et al., 2024).

Due to the complexity and the number of factors that can influence audience attention in VR, the spatial cinematographer must reduce those factors and plan for principal photography accordingly. Choosing the FOV is an important part of that, and while we have touched on 360° and 180° previously as popular formats, there are other possible formats that the spatial cinematographer can potentially use, such as:

- 200° (Closest to 180° but allows for more head movement)
- 220° (Allows for slightly more head movement)
- 240° (A FOV closer to 360° but still not as wide)
- 63° (slightly larger than the FOV of our eyes)

All these formats deviate from the intended/designed use cases for the available VR camera systems. But it is always important to remember that viewer attention is often focused on the centre of the frame (He and Liu, 2024).

Returning to it, however, as we have previously seen, narrower FOVs are a lot easier to direct audience attention around, and this is why CVR has transitioned primarily to being focused on VR180°. Having a narrower FOV than 360° allows the production quality to increase as the shots can be lit properly using more traditional cinema techniques; better production design is possible because a certain area is not being filmed due to the narrower FOV.

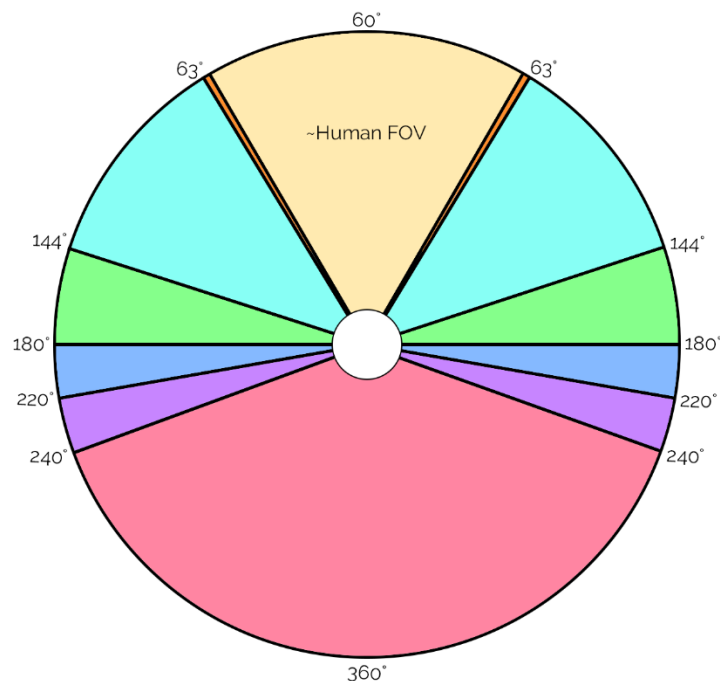


Figure 20: Top-down illustration of FOVs for VR. The viewer is in the centre.

Now, let us look at some illustrative examples from my practice to understand further how much of the frame is visible to the viewer in the HMD. The first example below is from the opening sequence of the first piece of my "Dreamstate" trilogy, of which the second part has already been mentioned in previous sections.

Below is an equirectangular screenshot in figure 20 from "Dreamstate" that represents the whole 360° space available to the viewer to look around.

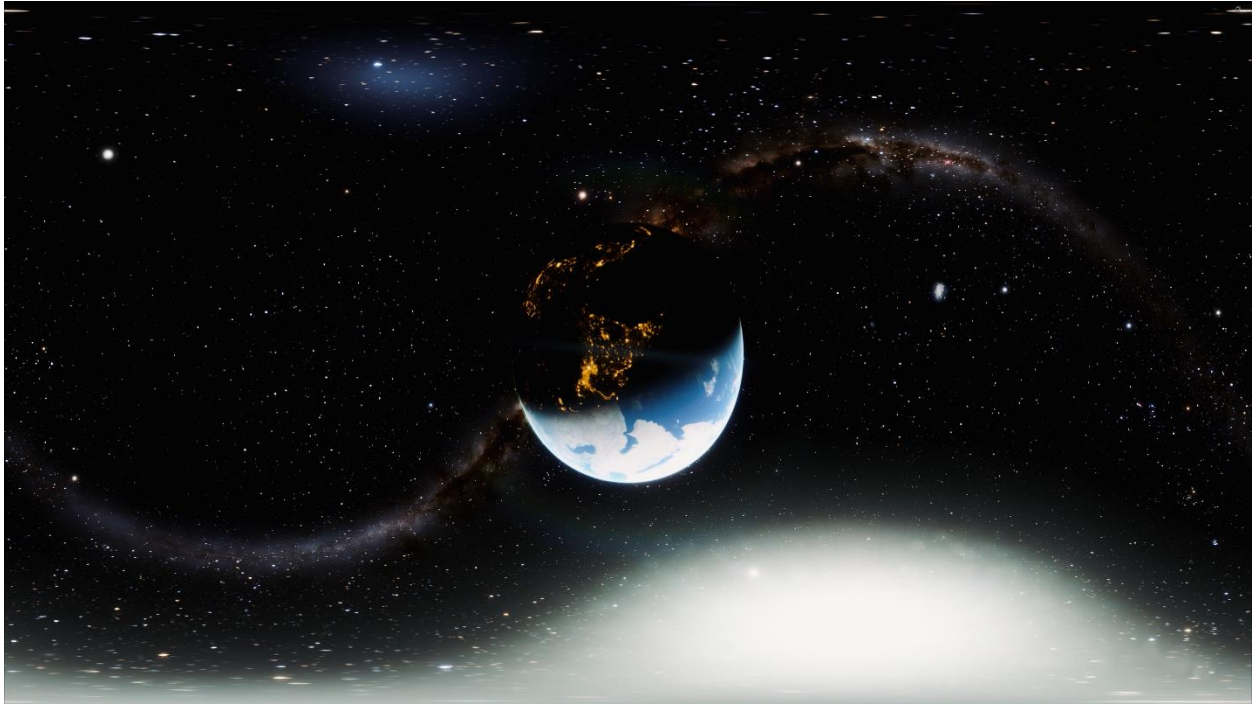


Figure 21: Opening sequence of "Dreamstate".

Figure 21 shows the various FOVs we discussed previously in this section, illustrated in a 16:9 aspect ratio, to give the reader an idea of how much of the frame will be available for the viewer to look around in.

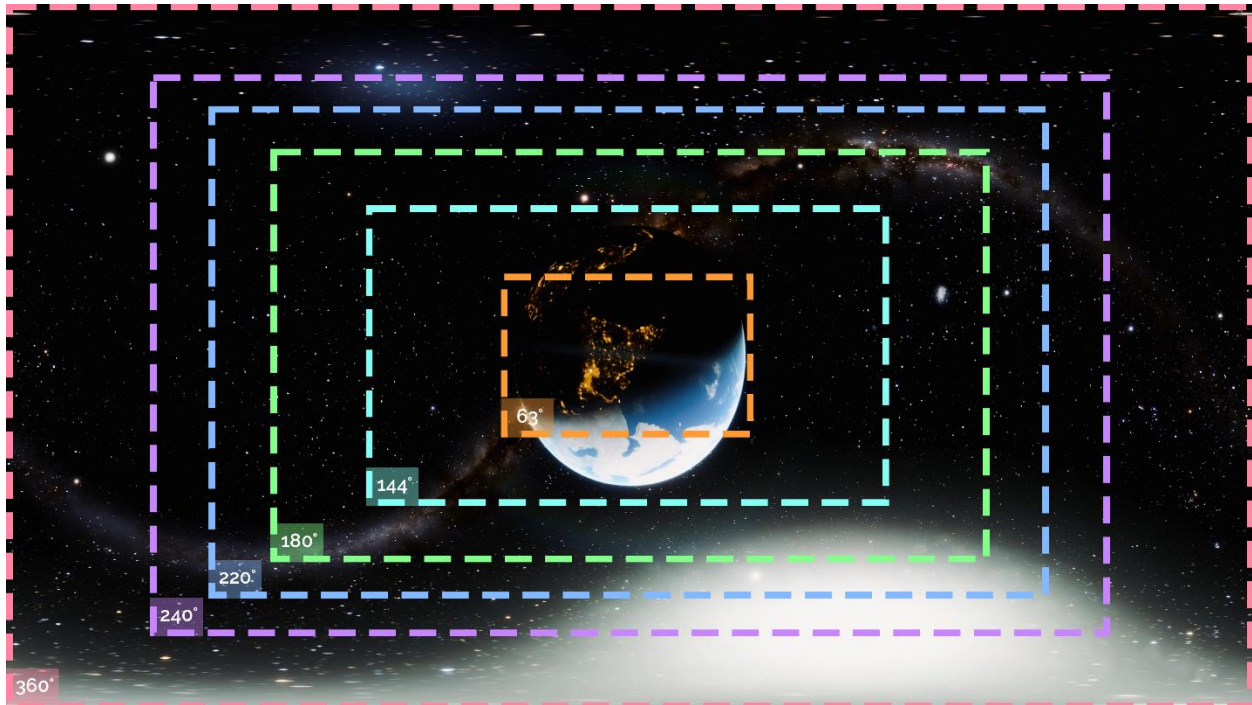


Figure 22: Opening sequence of "Dreamstate" showing FOV cropping.

Let us look at another example. Below in Figure 23 is a still photograph shot with Canon's 5.2mm Dual Fisheye lens, which captures an approximate 190-degree FOV. The previous example was a 2D CG VR 360° experience, and the one below is a 3D VR 180° still image.



Figure 23: ~190-degree FOV 3D still image without lens mask (SOOC).

Now, we will look at the image with a lens mask applied and how much of the frame gets cropped as we move to a narrower FOV. For the sake of simplicity, only the left eye will be illustrated.

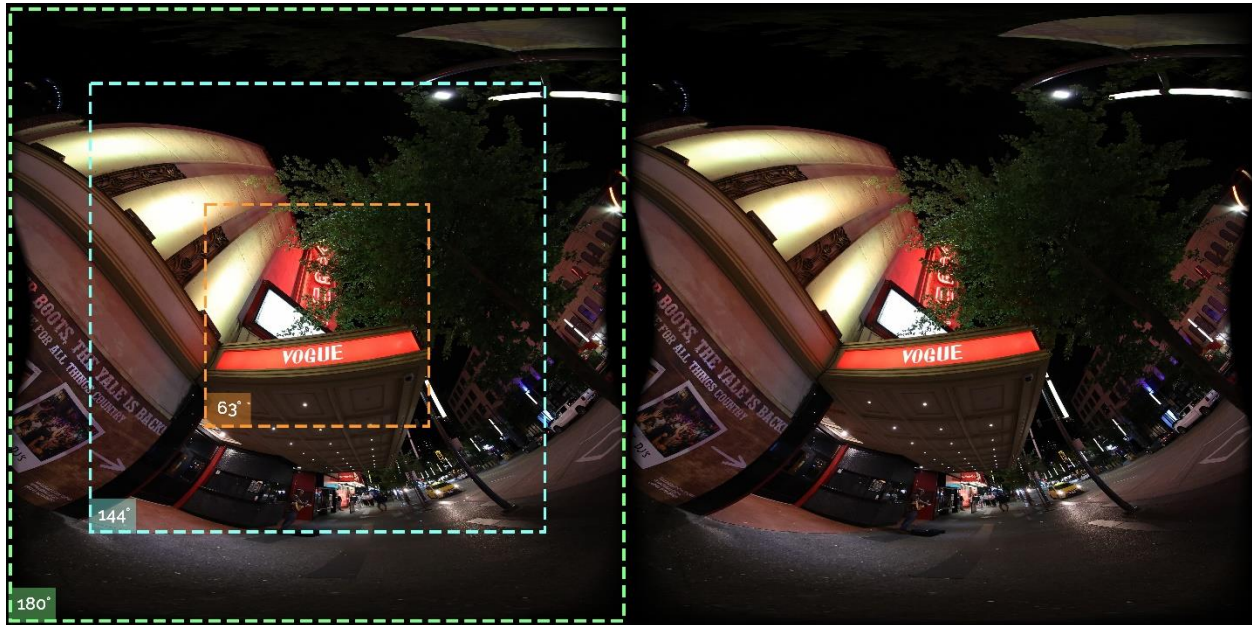


Figure 24: ~190-degree FOV 3D still image with lens mask via EOS VR utility & FOV crop illustrated.

In Figure 24 above, we see an approximation of how much space the viewer might have if we crop into a narrower FOV or shoot at that FOV in the first place.

We also see that the frame is automatically cropped to 180 degrees by applying the lens mask. By shooting at a *slightly* larger FOV of 190 degrees and applying a lens mask later, we ensure we always have a full 180-degree frame. By increasing the size of the mask, we can crop into the footage to make narrower FOVS; the same goes when working with 360°.

In conclusion, FOV is a powerful tool for the Spatial Cinematographer. The ability to manipulate the frame can direct the viewers' eyes to a point of interest, or it may create some emphasis, sense of constriction, or sense of expansion based on how the lens mask is being manipulated. It is a crucial tool that any spatial cinematographer must be aware of while planning during pre-production, and all these ideas can be prototyped using the aforementioned pre-vis techniques.

4.2 Choosing a Camera System

As CVR and spatial cinematography continue to evolve, a range of camera systems has become available, each catering to different levels of production quality, budget constraints, and usability requirements. This section categorizes VR camera systems into three primary tiers: consumer-level, prosumer-level, and professional-level setups. Each category offers distinct advantages and limitations, which the spatial cinematographer must carefully consider when selecting a system.

Consumer-Level VR Cameras

Consumer-level VR cameras are designed for entry-level users and casual content creators. These cameras typically offer ease of use, compact form factors, and affordability, making them suitable for hobbyists or those experimenting with VR content creation.

Notable consumer-grade VR cameras include:

- **Insta360 EVO:** Previously a popular choice, newer models have now surpassed this camera due to its outdated resolution and firmware limitations.
- **Kandao Qoocam 8K, EGO and Ultra 3:** The Qoocam 8K is a lightweight, portable 8K 360-degree camera. The EGO is a 3D camera with a FOV of around 66 degrees, and the Qoocam Ultra 3 is like the 8K but with more features. While Kandao's offerings provide impressive image quality for their price point, they suffer from software stability issues.⁸
- **Insta360 X-Series (X4, X3, etc.):** Positioned as 360-degree "action" cameras, these models offer solid performance compared to Kandao's offerings at a similar price point.⁹
- **CALF x VISINSE:** This is positioned as a more premium consumer camera. While promising, its firmware has presented operational challenges.¹⁰

⁸ See Kandao online product page at <https://www.kandaovr.com/consumer>

⁹ See Insta360 product page here:

https://store.insta360.com/consumer?i_source=website&i_medium=menu_button&i_campaign=consumer

¹⁰ See the CALF Product page here: <https://calfglobal.com/products/calf-visinse-3d-vr180-camera>

While these cameras provide accessibility, many lack professional-grade features such as LOG or RAW recording, advanced media storage options, robust power solutions, and reliable overheating management. These shortcomings make them less suitable for professional productions requiring extensive post-production workflows and higher image fidelity.

Prosumer-Level VR Cameras

The prosumer category bridges the gap between consumer-grade and high-end professional setups, offering improved image quality, better control over recording settings, and compatibility with professional editing software.

A significant advancement in this category is the **Canon EOS VR System**, which provides a range of VR lens options:

- **Canon RF 5.2mm f/2.8 Dual Fisheye Lens:** Designed for full-frame sensors, capturing a 190-degree field of view.¹¹
- **Canon RF-S 3.9mm F3.5 STM Dual Fisheye Lens:** Optimized for APS-C sensors, offering a 144-degree field of view.¹²
- **Canon RF-S 7.8mm F4 STM DUAL Lens:** A more recent addition, covering approximately 66 degrees, making it suitable for close-up and macro applications.¹³

¹¹ https://www.canon.ca/en/product?name=RF5.2mm_F2.8_L_Dual_Fisheye&category=/en/products/Lenses/VR-Lenses

¹² https://www.canon.ca/en/product?name=RF-S3.9mm_F3.5_STM_DUAL_FISHEYE&category=/en/products/Lenses/VR-Lenses

¹³ https://www.canon.ca/en/product?name=RF-S7.8mm_F4_STM_DUAL&category=/en/products/Lenses/VR-Lenses



Figure 25: Image via Canon Canada.¹⁴

The full-frame lenses are primarily used with Canon's **R5C, R5 Mark II, C80, and C400** cameras, and the APS-C lenses are currently only compatible with the Canon **R7** (Fensome and Kendrick, 2024). The Canon R5C and R5 MKII record up to DCI 8k at 60fps, the C80 up to DCI 6k at 30fps, and the C400 up to DCI 6k at 60fps, all in CRAW format. Meanwhile, the R7 records up to 4k Fine and up to 30fps.

CVR's ability to look "good" inside a headset is at a resolution of 8k, which translates to 4k per eye, and this is generally considered the baseline for CVR content. Lower resolutions lack detail and do not look as good inside a headset as more artifacts, and a general overall lower visual fidelity makes it just that harder for the viewer to get immersed.

When paired with the **Canon EOS VR Utility**¹⁵ and its Premiere Pro plugin, filmmakers benefit from a streamlined post-production workflow. The EOS VR utility handles all the stitching. It allows for one-click stabilization, Lens mask creation, and horizon correction, reducing the need for third-party software such as Mistika VR.

¹⁴ See Canon's EOS VR lens lineup here <https://www.canon.ca/en/products/Lenses/VR-Lenses>

¹⁵ Read more about the EOS VR utility more here <https://app.ssw.imaging-saas.canon/app/en/vru.html>



Figure 26: EOS VR Utility featuring some footage I shot in Venice during January 2025.

Professional-Level VR Cameras

Professional VR camera systems offer the highest quality, flexibility, and post-production control for high-budget productions and advanced immersive filmmaking. However, they also require post-processing software such as Mistika VR or Davinci Resolve with the KartaVR Plugin pack to perform tasks such as stitching, stereo correction, etc.

The most notable recent entry in this category is the **Blackmagic URSA Cine Immersive**, which features:

- **~17K overall resolution (8K per eye)**, the highest available in commercial VR filmmaking.
- Superior dynamic range and BRAW recording capabilities.
- Robust sensor performance, making it ideal for, large-scale productions.
- Native Integrated workflow with Davinci Resolve Studio.¹⁶

¹⁶ For more info, see <https://www.blackmagicdesign.com/ca/media/release/20240611-02>.

Other notable high-end immersive cameras include:

- **RED V-Raptor X** paired with Canon's dual fisheye lenses is the go-to for Cinematic VR production.¹⁷
 - 4k Per eye, 8K 120 FPS recording capabilities, and REDRAW Codec.
- **Apple's in-development in-house VR camera system**, details of which remain largely undisclosed due to confidentiality.¹⁸
 - We know it is in a similar spec range to the Ursa Cine Immersive at 8k per eye.

Selecting the appropriate VR camera system depends on the specific needs of a project, including budget, resolution requirements, workflow compatibility, and production scale. Consumer cameras provide accessible entry points but lack professional-grade features. Prosumer systems, particularly Canon's EOS VR lineup, offer a compelling balance between quality and affordability. Meanwhile, professional VR cameras like the Blackmagic URSA Cine Immersive and RED V-Raptor X push the boundaries of immersive filmmaking, albeit at significantly higher costs. As VR technology advances, the landscape of available camera options will evolve, offering new opportunities for filmmakers to innovate within immersive space.

¹⁷ See <https://www.youtube.com/watch?v=L7iP4zRD3cY> for more information.

¹⁸ Read more here <https://appleinsider.com/articles/24/10/11/apples-secretive-3d-cinema-camera-resurfaces-for-submerged>

Chapter 5: Post-Production

5.1 Image Stabilization and Upscaling

Image stabilization plays a critical role in CVR due to the unique challenges associated with camera movement. Unlike traditional filmmaking, where the camera's movement does not directly impact the viewer's sense of presence, the spatial cinematographer must carefully synchronize between visual motion and the viewer's physical orientation. Any discrepancy between these can lead to discomfort, including headaches and motion sickness (Li et al., 2021).

Even things such as the camera's position relative to the viewer's height can affect how the user experiences the scene (Rothe et al., 2019). The spatial cinematographer is heavily responsible for achieving smooth and stable imagery, paramount in VR filmmaking.

Spatial cinematographers will rely on meticulous shot planning and stabilization techniques to mitigate motion-related issues. The most common approach is to use locked-off shots, positioning the camera on a tripod or monopod to minimize unintended motion. Additionally, gimbals and Steadicam systems are widely employed to achieve fluid movement. Handheld shooting is strongly discouraged, as it introduces micro-jitters that can be challenging to remove in post-production, often exacerbating motion sickness in VR headsets.

There are three primary techniques for image stabilization in VR filmmaking:

1. **Physical Camera Stabilization:** This involves using hardware-based stabilizers, such as gimbals from manufacturers like DJI and Zhiyun. Professional systems like Steadicam, Flycam, or ARRI Trinity offer advanced stabilization capabilities for higher-budget productions.



Figure 27: *Myself with my Canon R7 w/ RF-S 3.9mm F3.5 Dual Fisheye Lens on my Zhiyun Crane 3 Lab Gimbal.*

2. **Post-Production AI Stabilization:** AI-driven software solutions provide an alternative for filmmakers who may not have access to high-end stabilization hardware. Topaz AI is widely regarded as the leading image stabilization and upscaling software.



Figure 28: Topaz Video AI 6 image stabilization – Self-shot footage from Venice in January 2025.

3. **Gyroscopic Stabilization:** Some VR camera systems, such as Canon's EOS VR system, incorporate gyroscopic data into the footage. When processed through the EOS VR Utility, this data enables precise image stabilization, which is considered an effective low-cost method for VR applications.



Figure 29: EOS VR Utility with image stabilization analysis in progress – Stability is prioritized for this specific clip – Self-shot footage from Venice in January 2025.

Despite AI-based stabilization's advantages, it presents certain challenges, particularly in stereoscopic VR. AI algorithms may introduce artifacts that result in inconsistencies between the left and right eye views, leading to an uncomfortable viewing experience in the headset.

The Role of Upscaling in VR Filmmaking

Upscaling has become a common practice in VR filmmaking due to the demanding technical requirements of high-resolution content. Shooting in native **8K 60fps RAW** is often impractical due to the need for high-performance camera systems, extensive storage capacity, and rapid data transfer rates. As a result, many filmmakers opt to shoot at lower resolutions, such as **4K**, **5.7K**, or **6K** and then upscale the footage using AI-enhanced software like **Topaz AI** to achieve **8K** (4K per eye) or even **16K** resolution (8K per eye) for premium headsets such as the Apple Vision Pro. Generally, **8K** (4K per eye) is considered the accepted minimum industry standard for VR.



Figure 30: Topaz Video AI 6 – Source footage is 3840x2160 (4k) at 30 fps i.e 2k per eye and is being upscaled to 15360 x 8640 (~16k) at 60 fps i.e 8k per eye – Self-shot footage from Venice in January 2025.

While upscaling can yield impressive results, its effectiveness depends heavily on the quality of the original footage. Poorly shot footage cannot be "saved" in post-production through upscaling, reinforcing the need for **skilled spatial cinematography**. Optimal lighting conditions are essential to ensure that stabilization or upscaling can be applied effectively.

Additionally, spatial cinematographers must account for the cropping effect inherent in image stabilization. Since stabilization algorithms compensate for motion by adjusting the frame, some peripheral details may be lost. This is not the case with AI-based image stabilization, but that tends to introduce artifacts. This is particularly significant in VR, where cropping can disrupt the viewer's sense of scale. Unlike traditional filmmaking, where a zoomed-in shot may be perceived as an intentional framing choice, excessive cropping in VR can distort the viewer's perspective. For instance, a minor miscalculation in stabilization may cause human subjects to appear disproportionately large, creating an unsettling experience where the viewer perceives themselves as significantly smaller than the on-screen characters. This can break immersion and negatively impact audience engagement.

Both image stabilization and upscaling are integral to achieving high-quality Cinematic VR experiences. While hardware stabilizers, gyroscopic correction, and AI-driven post-production techniques offer viable solutions, each approach presents its own limitations.

Likewise, upscaling can circumvent the challenges of high-resolution native capture, but it is not a substitute for well-executed cinematography. Ultimately, careful planning in stabilization and image resolution is necessary to maintain immersion and ensure an optimal viewing experience in VR.

5.2 Editing and Compositing

CVR presents unique challenges distinct from traditional filmmaking, particularly in editing. Unlike conventional cinema, Spatial Cinematography frequently involves multiple camera lenses or video sources that must be seamlessly stitched together into a unified stream. Once a labor-intensive manual task, this process has been significantly streamlined by camera systems and software solutions advancements.

Stitching and Initial Processing

Modern VR cameras, such as Insta360 and Kandao cameras, now offer in-camera stitching or proprietary software solutions that simplify the process. For instance, footage captured with these systems can be processed by dragging the files into the respective software and executing an automated stitching command with a few steps (Busch, 2024).

Similarly, Canon's EOS VR system utilizes a unique lens configuration that produces a side-by-side spherical image. This footage can then be converted into an equirectangular projection using the EOS VR Utility (Fensome and Kendrick, 2024), ensuring compatibility with non-linear editing (NLE) systems like Adobe Premiere Pro and DaVinci Resolve.

Stereo Correction and Compositing

Stereo correction is a critical step in VR post-production (J. Tong et al., 2022). While built-in correction tools, such as Canon's parallax and horizon correction features, can manage most adjustments, specialized software like Mistika VR provides more granular control over stereo compositing. Mistika VR allows editors to fine-tune depth alignment, ensuring a comfortable viewing experience (Terzić and Hansard, 2016). Once these adjustments are made, the footage is imported back into the NLE for further editing and sequencing.

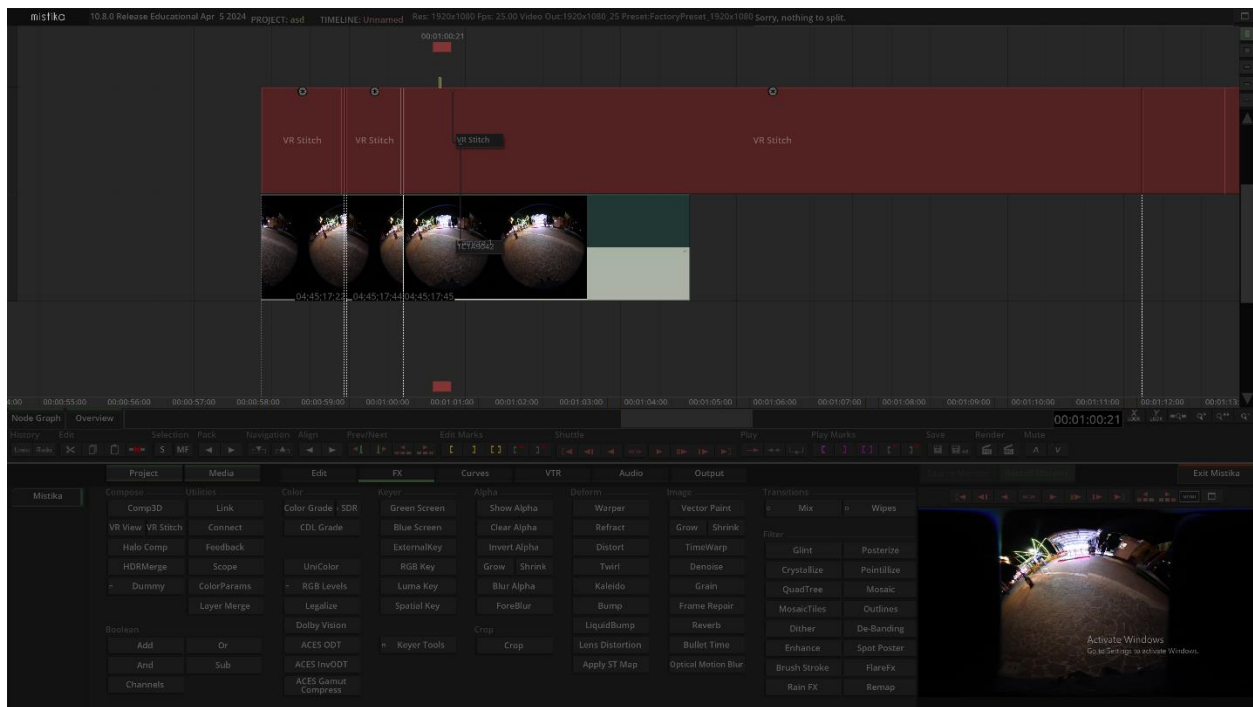


Figure 31: Example stitch of Canon EOS VR dual fish eye footage in Mistika Boutique – The spherical SBS footage can be seen in the timeline, and the stereo-composed equirectangular convert in the bottom right.

VR Editing in NLEs

Editing VR footage requires specialized tools for accurate previewing. Premiere Pro, for example, offers an in-headset live preview feature that enables editors to review footage in real-time, reducing the need for repeated exports and revisions. While editing on a flat screen remains an option, real-time VR previewing is recommended as it saves time and ensures precise spatial alignment. In DaVinci Resolve, a similar workflow can be achieved using the open-source KartaVR plugin (Hazelden, 2024), which utilizes Fusion to unwrap footage for headset viewing. Additionally, with the upcoming Ursa Cine Immersive, Blackmagic will enhance DaVinci Resolve Studio, transforming it into a fully native immersive video editing platform. However, the extent of its compatibility with third-party camera systems remains uncertain.

Metadata Injection and Export

Once the editing process is complete, the final step involves exporting and formatting the VR video with appropriate metadata. FFmpeg, a long-standing and widely used software tool, facilitates metadata injection to define critical attributes such as resolution, screen size, and stereo format (full SBS or half SBS). Specific metadata configurations are required for

different platforms. For instance, Apple Spatial Video (Swanson, 2024) necessitates specific metadata settings, while YouTube requires tags that identify the video format (e.g., VR180° or VR360) to ensure proper playback.

Considerations for VR Editing

One of the fundamental differences between traditional and VR editing is the necessity for the editor to understand stereo depth budget and alignment. While spatial cinematographers aim for optimal in-camera alignment, subtle stereo discrepancies can still arise. Even minor misalignments can cause viewer discomfort, (J. Tong et al., 2022) making it essential for editors to make precise stereo corrections during post-production.

Conclusion

Editing and compositing in Cinematic VR requires a combination of advanced software tools and a deep understanding of spatial cinematography principles. While technological advancements have simplified many aspects of the workflow, meticulous stereo correction, real-time VR previewing, and proper metadata handling remain critical for producing high-quality immersive experiences. As VR filmmaking evolves, integrating specialized editing tools within NLEs will further streamline the post-production process, enhancing Cinematic VR content's overall efficiency and quality.

Research Output: The Spatial Cinematography Framework Poster

To make these research findings more legible to a broader audience, I decided to take the key recommendations of my research and summarize them into a single graphical image. The advantage of this is that it is legible to a broader audience and more straightforward to disseminate than the format of my thesis document. The intention is to spread my research insights online to improve the field of spatial filmmaking today.

The poster has three sections: 1. Pre-production, 2. Production, and 3. Post-production.

Pre-Production:

- **Clarify Narrative Approach (3DOF vs. 6DOF):**
Clearly define whether the VR experience is immersive (non-interactive, 3DOF) or interactive (6DOF) to effectively guide production and storytelling decisions.
- **Conduct Thorough Pre-Visualization:**
Use advanced pre-visualization methods (e.g., Unreal Engine, NeRF scanning) to accurately simulate scenes, identify potential viewer attention issues, and streamline the shooting process.
- **Select Optimal Field of View (FOV):**
Determine the most appropriate FOV (180°, 360°, or custom) based on narrative requirements, directing attention, practicality in lighting, and intended viewer immersion.
- **Choose the Right Camera System:**
Carefully evaluate available camera systems (consumer, prosumer, professional) considering production scale, budget, resolution, and post-production workflow compatibility.

Production:

- **Apply Cinematic VR Framing Techniques:**

Strategically frame shots to guide viewer attention using spatial design, subtle gestural cues, and careful camera placement to maintain narrative clarity and immersion.

- **Ensure Stereo Comfort and Depth Budget:**

Meticulously calibrate stereoscopic 3D setups, maintain proper interocular distances, and consistently verify comfortable viewing to prevent viewer fatigue or discomfort.

- **Prioritize Stable Camera Movements:**

During shooting, use physical stabilization methods (tripods, monopods, gimbals) to minimize unwanted camera movement and viewer discomfort.

Post-Production:

- **Perform Precise Image Stabilization:**

Apply effective stabilization techniques (gyroscopic data, AI-driven software like Topaz Video AI) to ensure smooth and comfortable viewing experiences.

- **Optimize Resolution through Upscaling:**

When shooting at lower resolutions, utilize AI-enhanced upscaling techniques to achieve target VR resolutions (8K minimum) for maximum visual fidelity and immersion.

- **Implement Rigorous VR-specific Editing Practices:**

Utilized specialized editing workflows in software like Premiere Pro or DaVinci Resolve, ensuring careful stereo correction, compositing accuracy, and thorough real-time VR headset reviews to finalize the immersive experience.

The Spatial Cinematography Production Checklist:

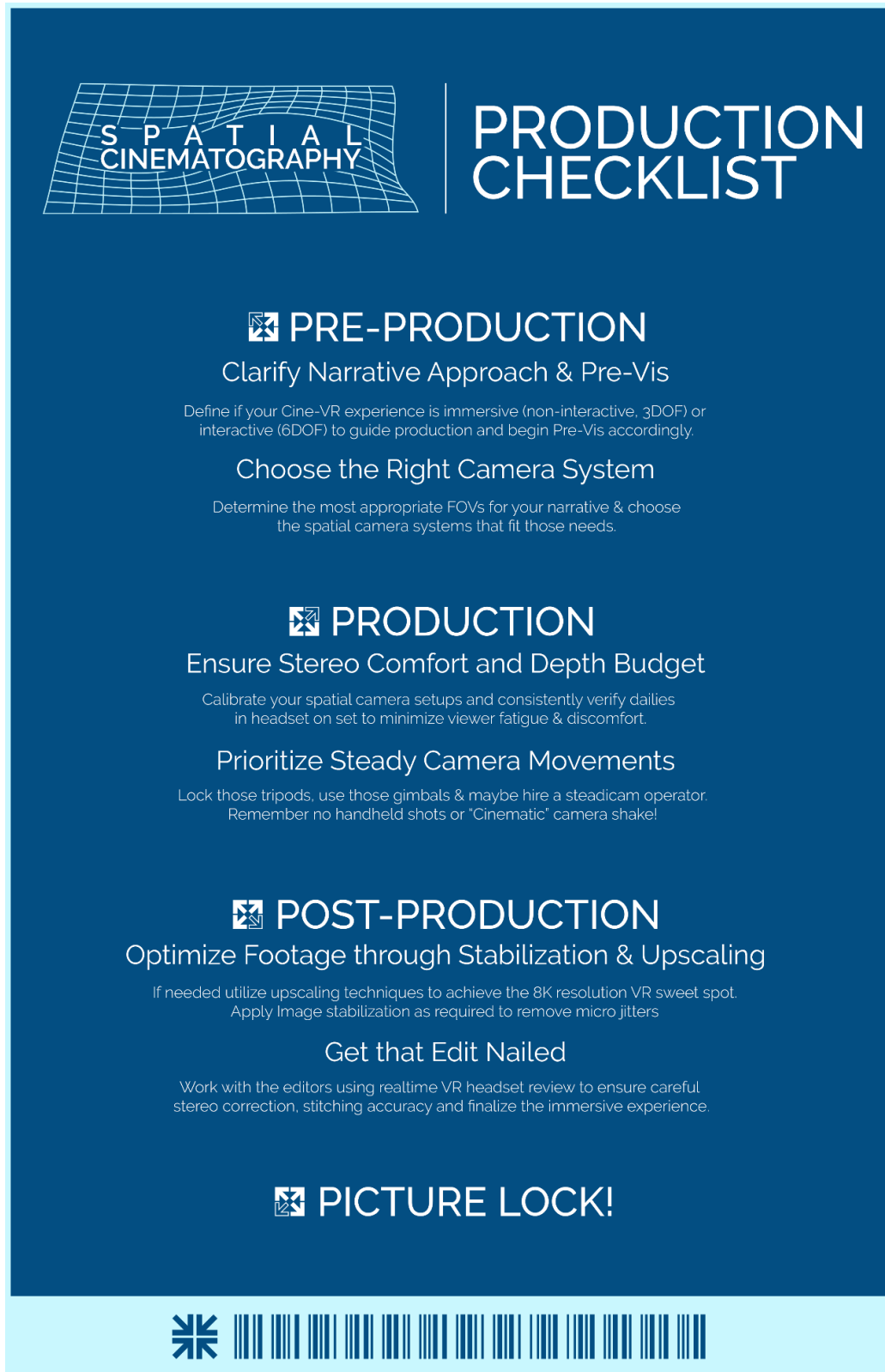


Figure 32: Spatial Cinematography Production Checklist.

The goals of this project reflect what experts suggest about making research findings legible to a broader audience. As Tufte notes, the page layout of this information matters: "Graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency" (Tufte, 2001). The goal here is to condense my findings into an easy-to-transport image that will be shared on Instagram, the web, and other digital platforms.

Conclusion and Future Directions

Spatial Cinematography represents more than a technical shift in filmmaking — it signals a deeper transformation in how stories are told, experienced, and remembered. This thesis has worked to map this new field by examining how spatial camera systems, non-interactive immersion, and production workflows reshape the language of Cinematic VR. By developing a practical framework, this research offers filmmakers a way to move beyond experimentation and begin crafting immersive works with greater precision, clarity, and emotional resonance.

At the heart of this project is a simple but profound idea: VR storytelling is not just about new tools — it demands new ways of thinking. As Marshall McLuhan famously argued, "the medium is the message" (McLuhan, 1964); the form of a medium shapes not just the content it carries but how it is understood. Spatial Cinematography embraces this reality, recognizing that VR is not simply another platform for cinema but a fundamentally different way of structuring narrative, space, and presence.

As hardware advances and technologies like AI-driven production and real-time rendering evolve, the challenge will not be merely to adapt, but to rethink the nature of narrative space itself. Spatial Cinematography provides a good foundation for this shift, helping to bridge the technical and creative divide that has held back the medium's full potential.

The accompanying framework and poster aim to make this knowledge accessible to a wider community of creators, opening doors for more diverse, sophisticated, and emotionally powerful VR films. As my upcoming Cine-VR work, *The Sound of One Eye Closing*¹⁹, moves toward its premiere at the Venice Biennale²⁰, it stands as proof that immersive cinema is not a passing trend — it is a medium still inventing its own grammar.

The future of VR storytelling will belong to those who see immersion not as a gimmick, but as a canvas. Spatial Cinematography invites filmmakers to claim that canvas — to

¹⁹ BCC press release: <https://www.labiennale.org/it/news/biennale-college-cinema-aperto-il-bando-italiano-novit%C3%A0-e-aggiornamenti>

²⁰ Read more about this year's BCC-I projects here
https://collegecinema.labiennale.org/en/prog_immersive_24/

experiment boldly, to question assumptions, and to create worlds that audiences can truly inhabit.

Bibliography:

Ackermann, E. (2018, June 15). *Introducing VR180 Creator, simplifying the video editing process.*

Google. <https://blog.google/products/google-ar-vr/introducing-vr180-creator-simple-video-editing/>

AMPAS. (2017, October 27). *THE ACADEMY'S BOARD OF GOVERNORS AWARDS AN OSCAR®*

TO ALEJANDRO G. INÁRRITU'S "CARNE Y ARENA" VIRTUAL REALITY INSTALLATION |

Oscars.org | Academy of Motion Picture Arts and Sciences.

<https://www.oscars.org/news/academys-board-governors-awards-oscarr-alejandro-g-inarritus-carne-y-arena-virtual-reality>

Apple Newsroom. (2023). *Apple introduces spatial video capture on iPhone 15 Pro.* Apple

Newsroom (Canada). <https://www.apple.com/ca/newsroom/2023/12/apple-introduces-spatial-video-capture-on-iphone-15-pro/>

Ardal, D., Alexandersson, S., Lempert, M., & Abelho Pereira, A. T. (2019). A Collaborative Previsualization Tool for Filmmaking in Virtual Reality. *Proceedings of the 16th ACM SIGGRAPH European Conference on Visual Media Production*, 1–10.

<https://doi.org/10.1145/3359998.3369404>

Arden, S. (2012). *Adventures in Stereo: Stereoscopic Cinema in the Age of a Digital Observer.*

<https://doi.org/10.35010/ecuad:2719>

Asch, T. (2022, January 16). *GoPro Dual HERO System Housing – An Extraordinary 3D Camera Rig – The Stereoscopy Blog.* <https://stereoscopy.blog/2022/01/16/gopro-dual-hero-system-housing-an-extraordinary-3d-camera-rig/>

Barnard, D. (2023, June 27). *Degrees of Freedom (DoF): 3-DoF vs 6-DoF for VR Headset*

Selection – VirtualSpeech. <https://virtualspeech.com/blog/degrees-of-freedom-vr>

Baoill, Andrew Ó. "Jenkins, H. (2006). *Convergence Culture: Where Old and New Media*

Collide. New York: New York University Press. 336 Pp. \$29.95 (Hardbound)." *Social*

- Science Computer Review*, December 3, 2007.
<https://doi.org/10.1177/0894439307306088>.
- Blackman, T., & Harley, D. (2024). Interpreting Apple's visions: Examining the spatiality of the Apple Vision Pro. *Platforms & Society*, 1, 29768624241283913.
<https://doi.org/10.1177/29768624241283913>
- Busch, A. (2024, March 20). *How to quickly stitch a 360 video with Insta360 Studio*. Mantis Sub Underwater Housings for Insta360 Pro / RS. <https://www.mantis-sub.com/academy/how-to-quickly-stitch-360-video-with-studio>
- Cannavò, A., Castiello, A., Praticò, F. G., Mazali, T., & Lamberti, F. (2024). Immersive movies: The effect of point of view on narrative engagement. *AI & SOCIETY*, 39(4), 1811–1825.
<https://doi.org/10.1007/s00146-022-01622-9>
- Carpio, R., Birt, J., & Baumann, O. (2023). Using case study analysis to develop heuristics to guide new filmmaking techniques in embodied virtual reality films. *Creative Industries Journal*, 1–22. <https://doi.org/10.1080/17510694.2023.2171336>
- Chatterjee, J., Spruyt, L., Pirson, N., & Vega, M. T. (2024). Effects of 6DoF Motion on Cybersickness in Interactive Virtual Reality. In L. T. De Paolis, P. Arpaia, & M. Sacco (Eds.), *Extended Reality* (pp. 21–37). Springer Nature Switzerland.
https://doi.org/10.1007/978-3-031-71713-0_2
- Cho, J., Kim, Y., Jung, S. H., Shin, H., & Kim, T. (2017). 78-4: Screen Door Effect Mitigation and Its Quantitative Evaluation in VR Display. *SID Symposium Digest of Technical Papers*, 48(1), 1154–1156. <https://doi.org/10.1002/sdtp.11847>
- Choi, H., & Nam, S. (2022). A Study on Attention Attracting Elements of 360-Degree Videos Based on VR Eye-Tracking System. *Multimodal Technologies and Interaction*, 6(7), Article 7. <https://doi.org/10.3390/mti6070054>

- Chulhyun, K. (2016, September 30). *A Comparative Study for Virtual Reality 360° Contents Shooting Equipments Based on Real World -Journal of Broadcast Engineering | Korea Science*. <https://koreascience.kr/article/JAKO201631267707579.page>
- De La Peña, N. (Director). (2012). *Hunger in Los Angeles*.
- Dixon, Steve. (2007). "Virtual Reality: The Search for Immersion.,".
<https://www.immersence.com/publications/2007/2007-SDixon.html>.
- Fensome, M., & Kendrick, M. (2024, November). *EOS VIRTUAL REALITY SYSTEM White Paper*. Canon Europe. <https://www.canon-europe.com/virtual-reality/>
- Gipson, J., Brown, L., Robbins, E., Gomez, J., Anderson, M., Velasquez, J., Ruiz, J., & Cooper, D. (2018). VR story production on Disney animation's "cycles." *ACM SIGGRAPH 2018 Talks*, 1–2. <https://doi.org/10.1145/3214745.3214818>
- Godard, J.-L. (Director). (2014). *Goodbye to Language (Adieu au langage)*.
- Hazelden, A. (2024, November 14). *Kartaverse | KartaVR*. Kartaverse.
<https://github.com/kartaverse/Kartaverse-Docs/>
- He, X., & Liu, Z. (2024). (PDF) A Novel Way of Estimating a User's Focus of Attention in a Virtual Environment. *ResearchGate*. https://doi.org/10.1007/978-3-319-91581-4_6
- IEEE. (2022). *The Differences between 3DoF and 6DoF, and Why—IEEE Digital Reality*.
<https://digitalreality.ieee.org/publications/degrees-of-freedom>
- IMAX Corporation. (1999). *The 15/70 Filmmaker's Manual: IMAX Corporation Website | PDF | 3 D Film | Camera*. Scribd. <https://www.scribd.com/document/635399609/Untitled>
- Iñárritu, A. G. (Director). (2017). *Carne y Arena*.
- Knorr, S., Kunter, M., Sikora, T., & Ide, K. (2024). (PDF) The Avoidance of Visual Discomfort and Basic Rules for Producing "Good 3D" Pictures. *ResearchGate*.
<https://doi.org/10.5594/j18236>

- Knorr, S., Ozcinar, C., Fearghail, C. O., & Smolic, A. (2018). Director's cut: A combined dataset for visual attention analysis in cinematic VR content. *Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production*, 1–10.
<https://doi.org/10.1145/3278471.3278472>
- Li, J., Reda, A., & Butz, A. (2021). Queasy Rider: How Head Movements Influence Motion Sickness in Passenger Use of Head-Mounted Displays. *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 28–38.
<https://doi.org/10.1145/3409118.3475137>
- Lynn, M. H., Luo, G., Tomasi, M., Pundlik, S., & E. Houston, K. (2020). Measuring Virtual Reality Headset Resolution and Field of View: Implications for Vision Care Applications. *Optometry and Vision Science*, 97(8), 573.
<https://doi.org/10.1097/OPX.0000000000001541>
- Mateer, J. (2017). Directing for Cinematic Virtual Reality: How the traditional film director's craft applies to immersive environments and notions of presence. *Journal of Media Practice*, 18(1), 14–25. <https://doi.org/10.1080/14682753.2017.1305838>
- Manovich, Lev. (2001). "The Language of New Media." MIT Press,
<https://mitpress.mit.edu/9780262632553/the-language-of-new-media/>.
- Milk, C. (Director). (2015). Clouds over Sidra. United Nations.
- MIXED. (2023, November 18). *The immersive turnaround: How VR180 3D cameras are conquering the (smartphone) market*. MIXED Reality News. <https://mixed-news.com/en/immersive-turnaround-vr180-3d-cameras-smartphone-market/>
- Morana, G. (2024). *Impact of Imaging and Distance Perception in VR Immersive Visual Experience*. <https://doi.org/10.18745/th.27468>

- Muender, T., Fröhlich, T., & Malaka, R. (2018). Empowering Creative People: Virtual Reality for Previsualization. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–6. <https://doi.org/10.1145/3170427.3188612>
- Murray, Janet H. (1997) "Hamlet on the Holodeck: The Future of Narrative in Cyberspace." MIT Press, <https://mitpress.mit.edu/9780262631877/hamlet-on-the-holodeck/>.
- Nielsen, Frank. (2005). "Surround Video: A Multihead Camera Approach." *The Visual Computer* 1, no. 1 : 92–103. <https://doi.org/10.1007/s00371-004-0273-z>.
- Park, S.-H. (2018). The Study on the Role of 3D Animated Pre-visualization in VFX FilmProduction. *Cartoon and Animation Studies*, 293–319. <https://doi.org/10.7230/KOSCAS.2018.51.293>
- Peddie, J. (2024, October 10). Remember Stereo 3D on the PC? Have You Ever Wondered What Happened to It? *ACM SIGGRAPH Blog*. <https://blog.siggraph.org/2024/10/stereo-3d-pc-history-decline.html/>
- PHI. (2017). *CARNE y ARENA (Virtually present, Physically invisible) | On Tour*. PHI. <https://phi.ca/en/carne-y-arena/>
- Rahman, A. R. (Director). (2022). *Le Musk*.
- Riggs, Stephanie. (2019) "The End of Storytelling: The Future of Narrative in the Storyplex." *End of Storytelling*. <https://www.endofstorytelling.com>.
- Rothe, S., Kegeles, B., & Hussmann, H. (2019). Camera Heights in Cinematic Virtual Reality: How Viewers Perceive Mismatches Between Camera and Eye Height. *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, 25–34. <https://doi.org/10.1145/3317697.3323362>
- Schleser, M., Remedios, D. J., Berrett, J., & Mathew, D. J. (2024). Imaginative storytelling – novel immersive production practices and processes for mobile cinematic,

- interactive 360-degree and real-time VR. *Media Practice & Education*, 1–19.
<https://doi.org/10.1080/25741136.2024.2426076>
- Swanson, M. (2024, March 7). *Spatial Video – Mike Swanson's Blog*.
<https://blog.mikeswanson.com/spatial-video/>
- Szita, K., & Gander, P. (2021). The Effects of Cinematic Virtual Reality on Viewing Experience and the Recollection of Narrative Elements. *PRESENCE: Virtual and Augmented Reality*.
https://www.academia.edu/107267704/The_Effects_of_Cinematic_Virtual_Reality_on_Viewing_Experience_and_the_Recollection_of_Narrative_Elements
- Terzić, K., & Hansard, M. (2016). Methods for reducing visual discomfort in stereoscopic 3D: A review. *Signal Processing: Image Communication*, 47, 402–416.
<https://doi.org/10.1016/j.image.2016.08.002>
- Tong, J., Wilcox, L. M., & Allison, R. S. (2022). The impacts of lens and stereo camera separation on perceived slant in Virtual Reality head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(11), 3759–3766. IEEE Transactions on Visualization and Computer Graphics.
<https://doi.org/10.1109/TVCG.2022.3203098>
- Tong, L., Lindeman, R. W., & Regenbrecht, H. (2021). Viewer's Role and Viewer Interaction in Cinematic Virtual Reality. *Computers*, 10(5), Article 5.
<https://doi.org/10.3390/computers10050066>
- Truong, A. (2015, December 28). *Virtual reality filmmakers are discovering new things about the power of audio*. Quartz. <https://qz.com/579185/virtual-reality-filmmakers-are-discovering-new-things-about-the-power-of-audio>
- Tufte, E. (2001). *The Visual Display of Quantitative Information*. Edward Tufte.
<https://www.edwardtufte.com/book/the-visual-display-of-quantitative-information/>

- Vosmeer, M., & Schouten, B. (2017). Project Orpheus A Research Study into 360° Cinematic VR. *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, 85–90. <https://doi.org/10.1145/3077548.3077559>
- Wallworth, L. (Director). (2016). *Collisions*.
- Williams, E. R., Love, C., Love, M., & Durado, A. (2021). Cine-VR: A new medium. In *Virtual Reality Cinema*. Routledge.
- Xu, L., Agrawal, V., Laney, W., Garcia, T., Bansal, A., Kim, C., Rota Bulò, S., Porzi, L., Kotschieder, P., Božič, A., Lin, D., Zollhöfer, M., & Richardt, C. (2023). VR-NeRF: High-Fidelity Virtualized Walkable Spaces. *SIGGRAPH Asia 2023 Conference Papers*, 1–12. <https://doi.org/10.1145/3610548.3618139>
- Zhang, Y., & Weber, I. (2023). Adapting, modifying and applying cinematography and editing concepts and techniques to cinematic virtual reality film production. *Media International Australia*, 186(1), 115–135. <https://doi.org/10.1177/1329878X211018476>