SMPTE 2025 MTS Manuscript template

Optimizing 180º Stereoscopic Cinematography for the Apple Vision Pro

Stanley Lin

University of Southern California

Nathan Fairchild

University of Southern California

Spencer Lin

University of Southern California

Akaash Tumuluri

University of Southern California

Jordan Halsey

University of Southern California

Scott Fisher

University of Southern California

Written for presentation at the

SMPTE 2025 Media and Technology Summit

Abstract. *The Apple Vision Pro establishes a new benchmark for image fidelity in head-mounted displays, creating both opportunities and challenges for cinematic VR filmmaking. Producing content that fully leverages its display capabilities requires reconsideration of image acquisition, post-production, and delivery workflows. This paper presents a replicable 180º orthostereoscopic imaging pipeline designed to balance accessibility, reliability, and cinematic quality while meeting the demands of next-generation virtual and mixed reality displays. Using our experimental VR short film* A Larger Living Space *as a case study, we outline an end-to-end workflow that integrates structured capture techniques, AI-based image enhancement, and scalable visual effects methods. Our pipeline emphasizes preservation of fine detail, reduction of noise and artifacts, and seamless integration of practical and virtual lighting within wide fields of view offered by dual-fisheye lenses available to the market. By uniting prosumer-grade tools with scalable post-production methods, our approach demonstrates that high-fidelity immersive video can be achieved outside of large-scale studio contexts. This work contributes a consolidated framework for immersive cinematography and provides a foundation for future exploration into higher bit depth capture, high frame rate imaging, and real-time virtual set integration for VR film productions.*

Keywords. Virtual Reality, Orthostereoscopic Imaging, Virtual Reality Film Production Pipeline, Cinematography, AI-Assisted Film Production

Fig. 1. Color graded frame from *A Larger Living Space*, showcasing the fully denoised and upscaled footage with a computer-generated ceiling seamlessly noise-matched to the live-action plate, 2025.

1 Introduction

The Apple Vision Pro (AVP) introduces an unprecedented standard of image fidelity in head-mounted displays (HMDs), enabling highly immersive VR cinematic experiences. However, leveraging its full potential requires rethinking both acquisition and post-production workflows for stereoscopic content. In this paper, we present a novel imaging and post-production pipeline that addresses the core challenges of 180º stereoscopic cinematography optimized for AVP delivery while using our original VR short film, *A Larger Living Space*, as a case study. Our approach integrates prosumer-grade camera systems with AI-enhanced post-processing techniques to achieve high-fidelity immersive video. This approach can be broken down into three stages: image acquisition, AI-assisted image enhancement, and VFX compositing.

We begin by outlining an optimized imaging pipeline built around the Canon R5 C and the RF 5.2mm f/2.8 Dual Fisheye lens, henceforth referred to as the Camera and the Lens, respectively. While this system offers a complete and accessible orthostereoscopic imaging camera package with a 190º angle of view, it also presents significant limitations including suboptimal sensor noise performance, low optical resolving power, and framing/lighting constraints inherent to ultra-wide angle-of-view (AOV) capture. To address these limitations, we implemented structured image acquisition strategies such as a one-stop overexposure method to improve sensor noise performance, aperture selection to balance resolving power and light transmission, and a hybrid lighting design that leveraged virtual ceiling extensions, embedded LED practicals, and rigged overhead space lights.

Crucially, we developed a scalable post-production pipeline designed to meet the display demands of the AVP. Footage was processed through an AI-enhanced workflow that included denoising and upscaling to 16K using Topaz AI, ensuring maximal clarity when downsampled on AVP’s 4K-per-eye displays. Our workflow preserved high-frequency details while minimizing motion discontinuity through 60p capture and judder-free encoding, significantly enhancing viewer comfort and realism. This represents a departure from legacy post-processing techniques like total variation and wavelet denoising, which overly degrade detail critical to the immersive experience.

We also developed a visual effects (VFX) and compositing pipeline using tools such as Houdini, 3D Equalizer, and Nuke, relying on Pixar USD for render-agnostic scene assembly. Our virtual set extensions (particularly for ceiling replacement) enabled sophisticated lighting design within the extreme AOV of the Lens without compromising the frame. All final renders were delivered in MV-HEVC with embedded spatial metadata to ensure stereo compatibility and HMD-optimized playback.

This research provides a blueprint for high-fidelity immersive film production using accessible equipment, and offers a comprehensive documented technical workflow specifically tuned for the Apple Vision Pro. Beyond technical enhancements, our imaging pipeline contributes meaningfully to immersive storytelling by supporting more cinematic lighting, refined viewer gaze control through luminance contrast, and more naturalistic motion presentation.

Our work aims to democratize high-quality VR filmmaking by making high-fidelity cinematic results attainable for lower budget productions. We hope to spark further exploration into 180º stereoscopic content design, enabling independent creators to meet emerging standards in XR media consumption. Future work will explore higher bit depth acquisition, 120p capture, and real-time motion-tracked VFX set extensions using camera telemetry data for accurate CG scene integration and single or multi-axis camera movements.

2 Inspirations and Related Works

Recently, there has been an increasing interest in stereoscopic VR capture systems, both in commercial and academic domains. Camera systems such as the KanDao QooCam EGO, Z CAM K1 Pro, and the upcoming Blackmagic URSA Cine Immersive each target immersive imaging with varying trade-offs in resolution, affordability, and workflow maturity. The Blackmagic URSA Cine Immersive, in particular, is positioned as the most complete high-end solution: bundled with DaVinci Resolve Studio, it offers a highly integrated camera-to-headset workflow. However, this functionality is only accessible through the purchase of the camera itself, which retails for $29,995 USD [Ben Allan 2025], effectively limiting its availability to creative teams with substantial budgets.

Prior academic work has largely focused on isolated aspects of immersive imaging rather than integrated workflows. Studies and articles have examined the relationship between frame rate and simulator sickness [Wang et al. 2023], evaluated noise reduction algorithms [Yang et al. 2018], and individual camera functionality benchmarking [Behiri 2023]. Similarly, industry documentation from manufacturers provides specifications and operational guidelines but rarely addresses how these devices can be systematically deployed to achieve cinematic imagery in VR contexts.

To verify this gap, we conducted a scoping review of immersive imaging literature between 2020–2025, searching multiple electronic databases with terms including stereoscopic video, immersive video, and VR cinematography. While we identified research on 360° video and volumetric capture pipelines, we found no studies that consolidated best practices into an end-to-end 180° orthostereoscopic imaging workflow suitable for modern head-mounted displays.

Our contribution addresses this absence. Unlike existing documentation, this work is specifically optimized for the Apple Vision Pro, whose 34 PPD acuity and high dynamic range (HDR) display imposes stricter sharpness and noise thresholds than traditional cinema. The result is a practical framework enabling filmmakers and creatives to achieve high-fidelity immersive imagery from prosumer stereoscopic hardware at a fraction of the cost of high-end systems like the URSA Cine Immersive coupled with its immersive media post-production suite.

3 Imaging Hardware and Production Challenges

In the context of VR cinematography, our camera system presents limitations stemming from the Camera's sensor and processing characteristics and the Lens's optical design. The most significant challenges encountered during production included suboptimal noise performance and limited dynamic range of the Camera, poor optical resolving power of the Lens, and the extreme AOV of the Lens, which introduced substantial framing and lighting constraints.

3.1 Camera System

We selected the Camera and Lens based on the following considerations:

- Native RF mount compatibility with full lens-to-body communication
- Full-frame image format and 8192 x 4320 (8K) sensor resolution, providing sufficient image detail for further AI resolution upscaling and denoising
- High frame rate (59.94p) support at full resolution and image format, ensuring smooth motion continuity for HMD playback
- 190º AOV, producing a high degree of VR immersion when viewing footage in HMDs

3.2 Sensor and Recording Limitations

The Camera’s 8K, 36 × 24 mm full-frame sensor provides an acquisition resolution sufficient for subsequent AI-based upscaling and denoising processes. When paired with the dual fisheye Lens, the single 8K sensor is divided into two circular fisheye images corresponding to left- and right-eye views, resulting in an effective image diameter of approximately 3684 pixels per eye [Canon Inc. 2025]. In comparison, the Apple Vision Pro (AVP) renders at 3660 × 3220 pixels per eye [David Heaney 2024] with a 100° AOV [Nivsus 2024], substantially narrower than the 190° capture AOV of our imaging system. Because the AVP displays only a 100° subset of the original 190° capture at any given moment, the effective resolution per displayed pixel is reduced relative to our full-frame acquisition. This spatial “cropping” effect necessitates further post-production enhancement, including AI-driven upscaling, to ensure that the displayed imagery fully exploits the native per-eye resolution of the AVP and maintains perceptual fidelity within the HMD.

High recording frame rates are essential to preserve motion continuity in VR, thereby reducing user discomfort and maintaining perceptual immersion. The AVP supports playback at frame rate multiples of 24p and 30p up to 100p, enabling judder-free rendering when source material is properly matched. Conventional cinematic frame rates such as 24p or 25p, while aesthetically established in traditional filmmaking, have been demonstrated to be inadequate in VR contexts, where sub-60p acquisition may introduce perceptible judder and simulator sickness (SS) [Wang et al. 2023]. For this reason, principal photography was conducted at the Camera's maximum supported frame rate of 59.94p (60p) at full sensor resolution, recorded in Canon 12-bit Cinema RAW Light.

Exploiting the full 8K 60p recording capabilities of the Camera requires use of Canon's 12-bit Cinema RAW Light format, which omits in-camera noise reduction. Consequently, RAW footage exhibited elevated sensor noise, particularly within shadow regions [Behiri 2023]. These limitations were compounded by the project's low-light, gritty apartment setting. In effort to mitigate these issues, Canon Log 3 encoding was selected for its lower noise floor, improved shadow handling, and enhanced color retention relative to other log profiles [Canon Europe 2023]. Nevertheless, even at the Camera's native ISO settings of 800 and 3200, the resulting footage displayed significant noise in dark regions and susceptibility to highlight clipping, as illustrated in the left panel of Figure 2.

Fig. 2. Side-by-side comparison of unprocessed RAW footage (left) and AI-denoised, upscaled footage (right), 2025.

3.3 Optical Characteristics and Limitations of the Lens

The Lens's 190° AOV and 60mm interaxial distance enable stereoscopic capture with an orthographic perspective, meaning subjects are rendered in a three-dimensional manner at a true-to-life scale with minimal perspective distortion with respect to the Camera's placement and perspective [Lipton 1982]. As the human interpupillary distance is between 63mm to 65mm, the optical characteristics of the Lens are very similar to our natural vision. This enhances the immersive realism of our viewing experience [Lipton 1982]. However, the ultra-wide AOV also introduces significant framing and lighting challenges, complicating light fixture placements and ambient fill control.

Optical deficiencies in the Lens further degraded image quality. Given the necessity of maximizing light transmission to combat sensor noise, our initial approach prioritized shooting at the lens's widest aperture (f/2.8). However, testing revealed significant degradation in resolving power at this aperture, leading to unacceptable reductions in image sharpness.

4 Image Acquisition Optimizations

To systematically address these challenges, we structured our image acquisition pipeline to improve three critical areas of capture: exposure, optics, and lighting.

4.1 Exposure and Noise Mitigation Strategies

To compensate for the Camera's noise performance and limited dynamic range, we established a recording configuration optimized for color retention and noise reduction:

- Recording Format: Canon 12-bit Cinema RAW Light
- Resolution: 8192 x 4320
- ISO: 3200 (base 3200)
- Color Profile: Canon Log 3
- Frame Rate: 59.94p

Furthermore, we implemented a one-stop overexposure strategy [Canon Europe 2023] to further enhance shadow detail retention, reduce perceptible noise, and protect highlights. This approach involved systematically exposing for one stop above our target exposure level while monitoring footage through a custom pull-down lookup table (LUT) designed to simulate the correct exposure and intended final look in real time.

We standardized all shots at f/5.6, exposing for f/8 via a 1-stop overexposure method. This approach enabled the sensor to capture increased data in shadow regions while protecting highlights, effectively reducing noise accumulation in underexposed areas.

4.2 Optimizing Optical Performance of the Lens

Careful aperture selection was a critical factor in balancing optical resolving power against light transmission. Empirical testing demonstrated the following:

- f/2.8 exhibited significant optical softness, making it unsuitable despite its increased light transmission.
- f/7.1 yielded the highest resolving power, albeit still exhibiting suboptimal sharpness, but at the cost of reduced light transmission.
- f/5.6 provided an optimal compromise, offering an acceptable level of sharpness while securing an additional 2/3 stop of light transmission compared to f/7.1.

4.3 Lighting and Framing Adaptations for Ultra-Wide AOV

The ultra-wide AOV of the Lens meant traditional lighting setups had to be moved out of frame or hidden. This increases the distance of fixtures to the area of illumination and/or requires the use of smaller, more discrete fixtures. This reduces the total quantity of light in the scene and to the sensor. Yet again, this increases sensor noise. To circumvent these limitations, we employed a hybrid practical lighting and computer-generated production design strategy.

Fig. 3. Frame grab illustrating the set constructed without a ceiling to facilitate overhead lighting, 2025

4.3.1 Set Design for Optimal Fixture Placement

Our apartment set was designed with doorways and structural corners, allowing the use of conventional off-axis lighting setups to create strong accents and hidden hair lights without compromising framing, as seen in Figure 3. This layered lighting approach enabled us to preserve the visual integrity of the scene while guiding viewer attention to portions of the 180° frame that matters most to the narrative using luminance contrast. Lighting subtly directs the gaze without compromising the viewer’s freedom of exploration. We balanced interactive storytelling with cinematic presentation, motivating the viewer to follow the narrative without overt guidance.

4.3.2 Virtual Set Extensions to Motivate Overhead Lighting

In preparation for virtual ceiling set extensions, we built our set on a soundstage without a ceiling, as seen in Figure 3. This allowed us to rig fixtures on the soundstage’s overhead lighting grid to provide ambient fill and hard spots into the set. A computer-generated virtual ceiling was composited into the set in post-production, allowing us to maintain the illusion of an enclosed environment. Our overhead lights were motivated by the house lights on the virtual ceiling, ensuring spatial consistency between practical and CG elements. This ceiling-less set design was the primary method we used to achieve our target exposure levels while hiding all of our large lighting fixtures.

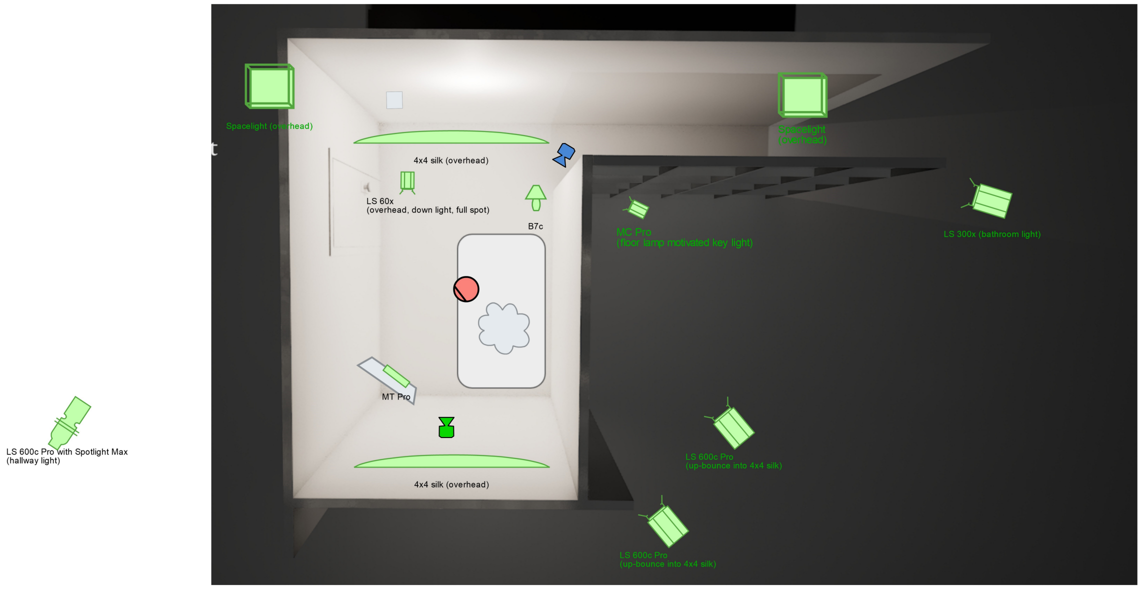

Fig. 4. Lighting diagram illustrating a set layout with lighting fixture, diffusion, and bounce placements, 2025.

4.3.3 Embedded Practical LED Fixtures

Discrete high-output and high-color-quality LED fixtures were hidden within practical set decoration pieces such as lamps and magnetized to walls out of frame to deliver subtle fill and hair lighting (refer to "B7C" and "MC Pro" in Figure 4), further enhancing depth and separation between subjects and the background.

5 Post-Production Workflow

While our acquisition pipeline does much to mitigate many of our imaging challenges, its most important role is to prime our footage for subsequent post-production. RAW footage out of camera still exhibits noticeable noise and optical resolution deficiencies. The final refinement of our footage into a polished VR film is ultimately achieved through our post-production pipeline.

5.1 Digital Intermediate Pipeline

When working with RAW camera footage recorded in its native format with correct gamma and gamut (in this case, Canon Log 3 with Canon Wide Gamut and monitored in P3), it is essential to distinguish between post-production workflows that incorporate computer-generated imagery (CGI) and those that do not.

For footage requiring no CGI integration, the pipeline employs HDR color grading in HDR Rec.2020, with final output targeted to the requisite color space of the destination device. For footage requiring CGI integration, the workflow transitions to an ACES-managed color pipeline within compositing, with transforms into HDR Rec.2020. An extensible color management design was prioritized to allow outputs to be transcoded easily into multiple formats and color spaces. Both workflows are illustrated in Figure 5.

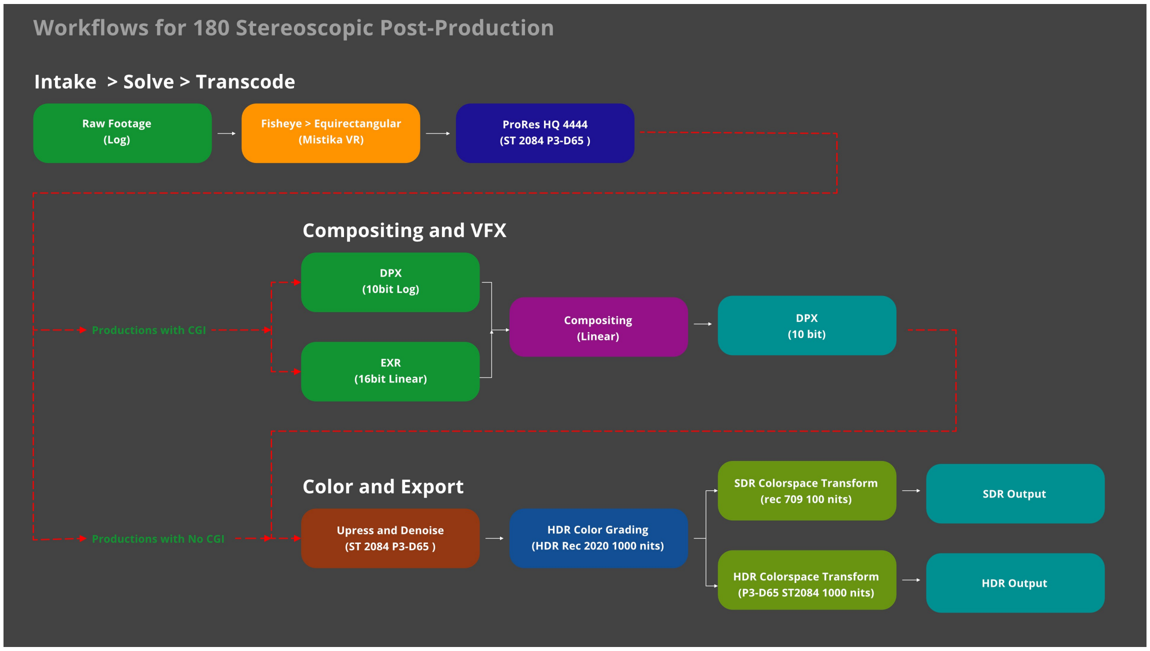

Fig. 5. Our digital intermediate pipeline flow chart detailing workflows for both CGI and non-CGI projects, 2025.

5.2 Strengths and Weaknesses of the Camera System for Our Post-Production Pipeline

Canon provides proprietary software for converting dual-fisheye imagery into stereoscopic 180° equirectangular format. However, in our testing, the software did not resolve imagery to an acceptable standard for research or production use. As a result, our workflow employed Mistika VR for this stage of the pipeline. Through camera-lens profiling conducted within Mistika VR, we identified two strengths and five primary weaknesses of our camera system:

- Camera system strengths:
  - Ability to shoot RAW at 60 frames per second
  - A total capture resolution of 8192x4320
- Camera system weaknesses:
  - Off-centered fisheye images projected from the lens onto the sensor, as seen in Figure 6
  - Low optical resolving power of the lens
  - Unknown exact IPD
  - Suboptimal noise performance, especially in low light scenarios
  - Focus inconsistency between the two fisheye projections, resulting from each lens group needing to be independently focused prior to each shot

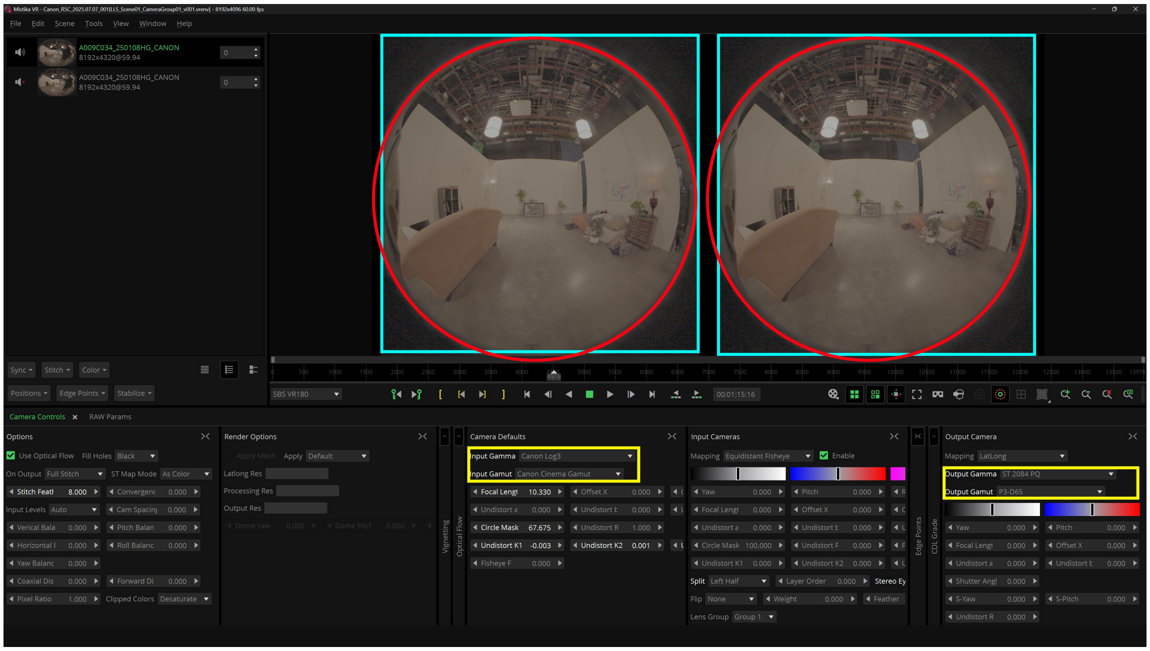

Fig. 6. A diagram illustrating off-center fisheye images resulting from the Camera and Lens, 2025.

5.3 Camera System Testing Methodology

Due to the extreme AOV and distortion characteristics of the Lens, a standard lens grid calibration shoot was insufficient. Using the software 3D Equalizer, we determined that a more rigorous approach involving a detailed series of tests supplemented with high-precision photogrammetry was required.

5.3.1 Photogrammetry

Conventional grid-based calibration proved inadequate for fisheye optics. To achieve higher accuracy, we employed 3D Equalizer in combination with high-precision photogrammetry. Custom marker sets were generated using RealityCapture, supported by a Leica Disto site survey. The resulting Leica Disto point-cloud data was integrated into the photogrammetric solution to ensure accurate global scale and orientation.

5.3.2 Footage Assets

To create a full lens profile for the Lens, we determined that we would shoot a series of tests to incorporate into 3D Equalizer. This would include a typical center grid lens shot with a static camera, a series of nodal camera pans for matchmoving, and a gimbaled handheld tracking shot. All of these would keep the main lens grid in full view.

5.3.3 Noise Testing

To evaluate noise performance, we tested multiple exposure strategies at both base ISO 800 and base ISO 3200. For each setting, exposures ranged from one stop under to two stops over. The best noise performance was ultimately observed at one and two stops overexposed.

5.3.4 Focus Testing

We shot a series of tests with incremental focus ranges starting at 9” and increasing a foot at a time with the goal of determining the best focal point for deep focus.

5.3.5 Footage Ingest, Transcode, Edit, and Output

Captured RAW footage was ingested through Mistika and split into left- and right-eye streams. Using Mistika’s lens guides, the off-centered fisheyes were corrected, and an initial geometric solve was calculated. Corrected footage was then output in Gamma P3-D65 and Gamut ST 2084 PQ. A P3 color space comparison can be observed in Figure 7.

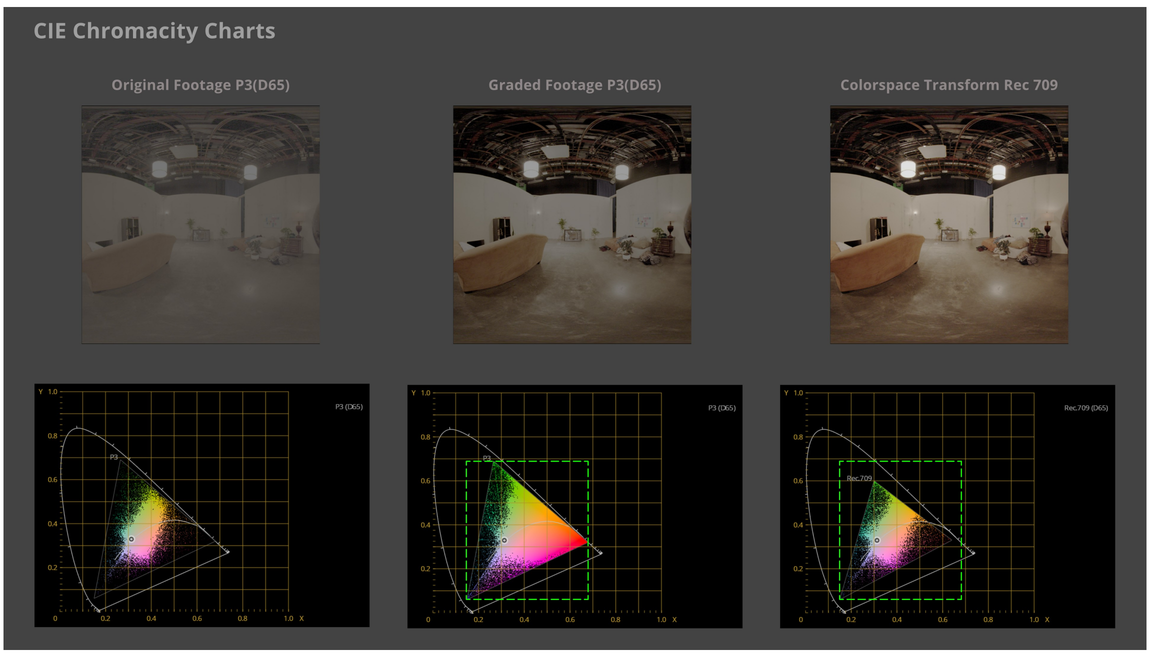

Fig. 7. Comparison between ungraded footage, graded footage within the P3 color space, and a color space transform to Rec 709 Gamma 2.4, 2025.

5.4 CGI Integration Testing

Plates were processed for compositing as 10-bit .DPX files. For scene alignment, we mapped the equirectangular plate to an inverted hemisphere oriented around the Z-Axis with five cameras locked at the zero position. We created five 90º angle rectilinear cameras oriented, +X,-X, +Y,-Y, +Z with a square filmback (24mm x 24mm).

5.4.1 Camera Validation – Testing Geometry Alignment

Scene alignment was validated using baseline measurements of camera height and on-set location data. Initial geometry was then aligned with the photogrammetry model.

To refine alignment, the base 3D geometry was rendered against the stereo backplates, shown in Figure 8. The comparison of stereo 180° equirectangular renders provided a feedback mechanism for iterative refinement of the geometry.

Once our 3D base model was refined, we then imported and oriented the 3D photogrammetry scan while making sure that there were never any scale adjustments, seen in Figure 9.

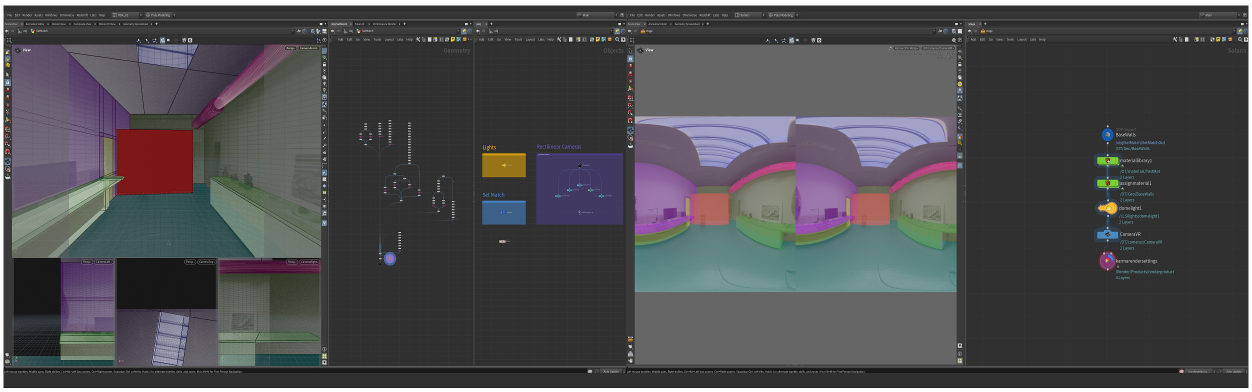

Fig. 8. 3D viewport used for alignment and the stereo renders over the backplate, 2025.

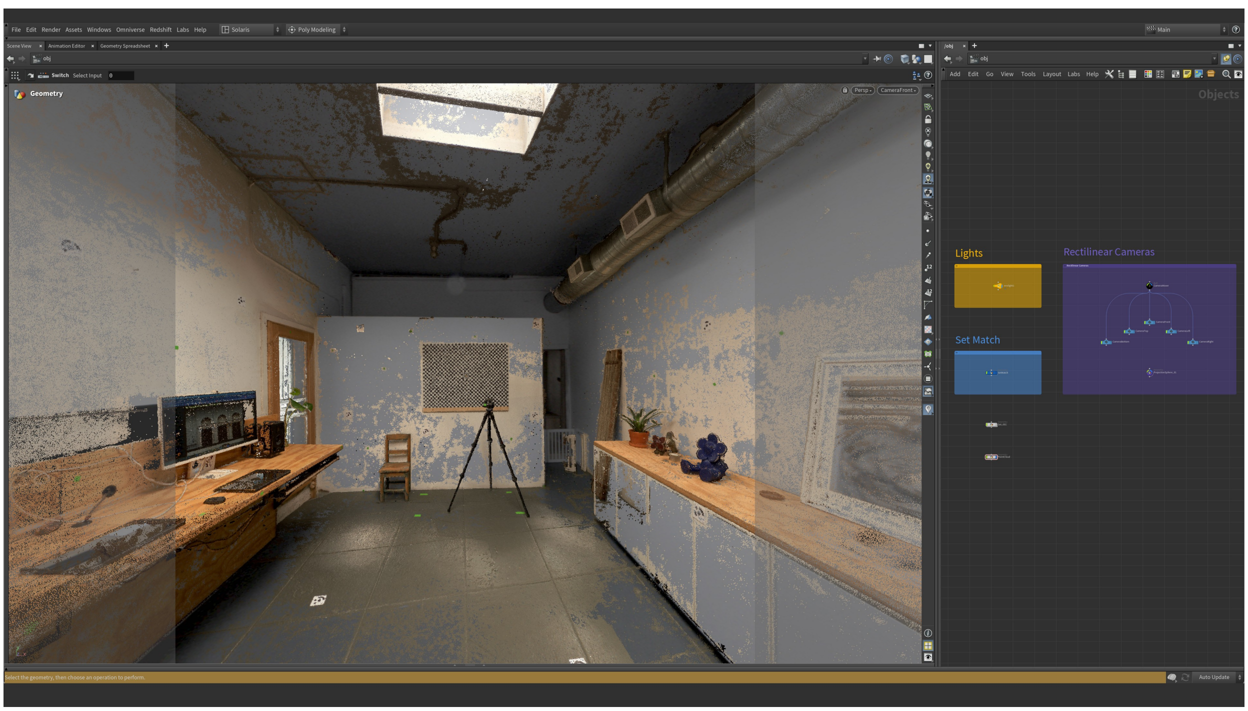

Fig. 9. 3D viewport showing the aligned photogrammetry with the backplate, 2025.

5.5 A Larger Living Space – Pipeline Integration Example

To ensure reproducibility of our physical environment in digital form, our practical apartment set was constructed using a modular flat system. All walls were fabricated from standardized 4x10 foot panels, which were subsequently replicated in a three-dimensional environment. These virtual modules served as a registration framework, aligned precisely with a base geometric model derived from on-set measurements. A custom Houdini Digital Asset (HDA) was developed to enforce consistent alignment across shots. The HDA also enabled the rendering of polylines over stereo equirectangular backplates, providing a mechanism to validate geometric correspondence and inter-pupillary distance (IPD) accuracy (alignment tool shown in Figure 10).

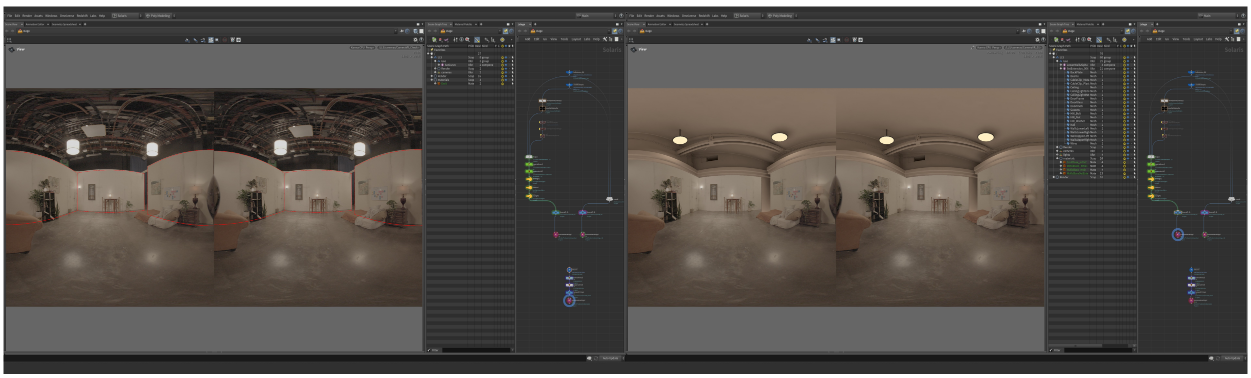

Fig. 10. Renders showing the alignment tool and the render over the backplate, 2025.

The base geometry was used as a structural guide for set extensions. These extensions were first rendered as a quality control pass against the original backplate and then employed as matte cut-outs to support compositing operations. An occlusion pass was generated in parallel to improve the perceptual blending of digital extensions with the live-action plate. Rendering of set extensions was performed within a Universal Scene Description (USD) pipeline, using Solaris and the Karma renderer with a MaterialX shading framework. The same USD-based workflow also facilitated the production of supplementary computer-generated imagery (CGI) in Unreal Engine, ensuring interoperability across platforms.

All 3D outputs were generated as multi-layer, linear EXR files, incorporating the following render passes: beauty, combined albedo, combined reflection, combined refraction, emission, position, normals, UVs, and a Fresnel component. These passes provided the necessary information to support advanced compositing operations. In order to preserve the raw fidelity of the renders, no denoising was applied during the rendering process.

Compositing was completed in Nuke with a procedural network, shown in Figure 11, that allows for work to be easily propagated across multiple shots. The 3D renders are grain and grade matched against the uncorrected footage and grain matched again in the compositing phase to ensure as close of a match as possible prior to the denoise and uprez phase.

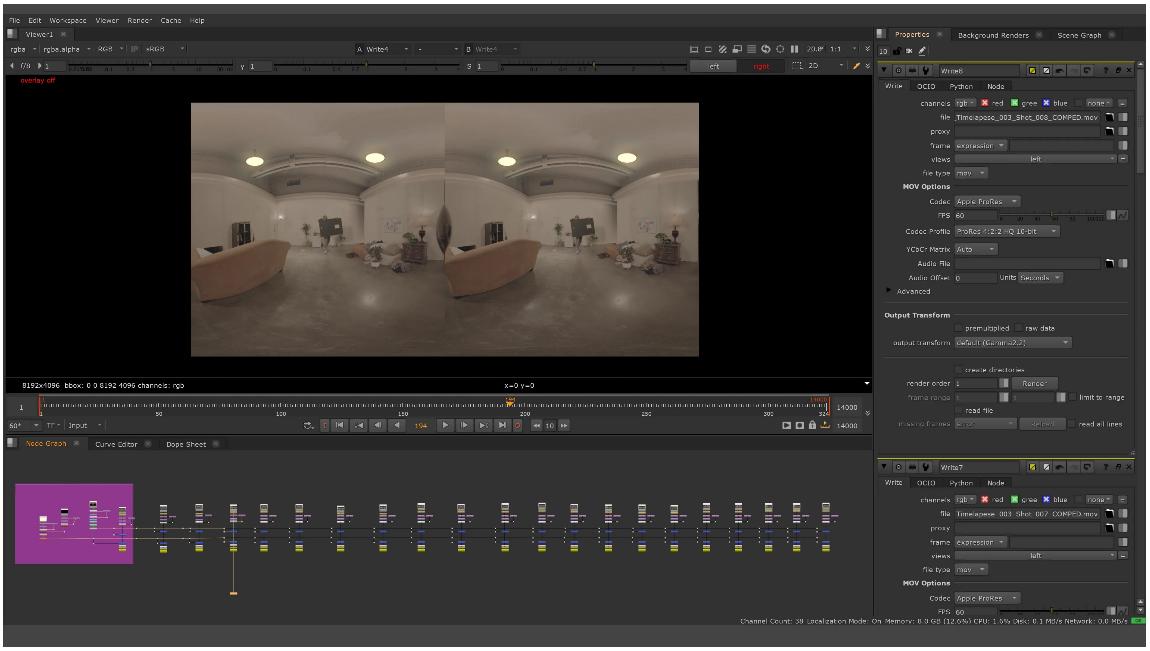

Fig. 11. Nuke compositing network, 2025.

5.6 AI Denoising Advantages

Traditional mathematical denoising techniques, such as Total Variation (TV) regularization and wavelet-based noise suppression, prioritize noise reduction at the cost of image sharpness [Yang et al. 2018]. While these methods can be effective for generalized noise control at viewing distances longer than that of an HMD, they often lead to texture smoothing and loss of high-frequency details. Any loss of edge sharpness or fine textures becomes immediately perceptible when viewing footage in the AVP. Given this, our denoising workflow must prioritize detail preservation over aggressive noise suppression, ensuring that high-frequency information remains intact while minimizing visible artifacts. Unlike mathematical denoising techniques that apply uniform processing across the frame, AI-based denoising dynamically adapts to edge contrast and texture complexity, allowing for strong noise suppression while retaining sharpness in high-detail areas [Tassano et al. 2019]. Empirically, this approach produces sharper imagery than traditional denoising methods. We selected Topaz Labs' Video Package for its ability to distinguish noise from true image structure, preventing over-smoothing.

5.7 Lens Profiling, Rendering, and Final Delivery

Initial lens profiling was completed using 3DEqualizer. For our static camera setups and 3D reconstruction, the left-eye equirectangular image is reprojected in SideFX Houdini, where it is converted from an equirectangular to a rectilinear projection for set extensions. Accurate set measurements and photogrammetry are critical for seamless CGI integration. Our 3D pipeline is render-agnostic and based on Pixar's Universal Scene Description (USD). Special attention is given to matching rendered noise with the original plate, and no denoising is applied to rendered imagery. Images are exported as layered .EXR files, containing at minimum depth world normals, Cryptomatte, and position passes. Compositing is completed in Nuke, and the final output is encoded in ProRes 4444 XQ before undergoing denoising and 16K upscaling in Topaz. The final master is encoded in MV-HEVC, a multi-view compressed format optimized for stereo output. Spatial metadata is embedded in the final delivery format. See Figure 1 for results.

6 Future Work

Significant work remains in advancing techniques and technologies for immersive cinematography. Future research can investigate imaging and processing approaches that extend dynamic range through higher bit-depth acquisition (e.g., 16-bit RAW workflows). Such methods would improve the ability to capture low-light scenes without relying on deliberate overexposure followed by post-production enhancements. As stereoscopic camera systems continue to evolve, further research into higher frame rate capture (≥120p) will be important for minimizing motion discontinuity and reducing simulator sickness. In addition, integrating motion-tracked VFX set extensions with externally recorded camera telemetry offers a pathway to more accurate alignment of computer-generated elements with physical lighting fixtures and camera movement.

7 Conclusion

The insights gained from *A Larger Living Space* informed the development of our capture-to-delivery workflow, which can be generalized across a range of stereoscopic cameras, head-mounted displays (HMDs), and post-production platforms. By combining established filmmaking practices with emerging VR and AI tools, we created a reproducible 180° stereo imaging pipeline designed for modern HMDs. Our acquisition approach (integrating deliberate overexposure, lensing trade-offs, and adapted lighting strategies) produced footage optimized for subsequent AI-driven denoising and 16K upscaling, allowing us to retain fine detail and minimize noise artifacts. Recording at higher frame rates further enhanced motion realism and viewer comfort. Overall, our findings suggest that immersive VR films can maintain the artistic intent and visual language of traditional cinema while meeting the technical requirements of next-generation HMDs, providing a practical foundation for both filmmakers and researchers advancing cinematic storytelling in VR.

References

Ben Allan. 2025. Blackmagic URSA Cine Immersive Camera– Measuring Resolution. Newsshooter. https://www.newsshooter.com/2025/07/04/blackmagicursa-cine immersive-camera-measuring-resolution/

Canon Europe. 2023. Enhance Filmmaking with Canon Log. Canon Europe. https://www.canon-europe.com/pro/stories/enhance-filmmaking-withcanon-log/

Canon Inc. 2025. Canon RF 5.2mm F2.8 L Dual Fisheye. Canon Canada. https://www.canon.ca/en/product?category=/en/products/Lenses/VR-Lenses& name=RF5.2mm_F2.8_L_Dual_Fisheye

David Heaney. 2024. Apple Vision Pro Extended Teardown Reveals Its True Active Resolution. Upload VR. https://www.uploadvr.com/apple-vision-proextended-teardown-reveals active-resolution/

Jialin Wang, Rongkai Shi, Wenxuan Zheng, Weijie Xie, Dominic Kao, and Hai-Ning Liang. 2023. Effect of Frame Rate on User Experience, Performance, and Simulator Sickness in Virtual Reality. IEEE Transactions on Visualization and Computer Graphics 29, 5 (2023), 2478–2488. doi:10.1109/TVCG.2023.3247057

Johnnie Behiri. 2023. Canon EOS R5 C Lab Test – Rolling Shutter, Dynamic Range and Latitude. CineD. https://www.cined.com/canon-eos-r5-c-lab-testrolling-shutter-dynamic range-and-latitude/

Lenny Lipton. 1982. Foundations of the Stereoscopic Cinema: A Study in Depth. Van Nostrand Reinhold, New York.

Matias Tassano, Julie Delon, and Thomas Veit. 2019. DVDNET: A Fast Network for Deep Video Denoising. In 2019 IEEE International Conference on Image Processing (ICIP). 1805 1809. doi:10.1109/ICIP.2019.8803136

Nivsus. 2024. Apple Vision Pro FOV: Limitations and the Quest for Enhancement. vskel. https://vskel.com/apple-vision-pro-fov/

Qi Yang, Cheolkon Jung, Qingtao Fu, and Hyoseob Song. 2018. Low Light Image Denoising Based on Poisson Noise Model and Weighted TV Regularization. In 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, Athens, Greece, 3199 3203. doi:10.1109/ICIP.2018.8451840

