# Notes from Apple's “Create immersive media experiences for visionOS” event

Anthony Maës Follow 43 min read · Nov 4, 2025



We'll also address the elephant in the room at the end.

Last week was a busy week for Acute Immersive. My co-founder Paul got married in India, and I flew back just in time to attend a two-day seminar at Apple’s developer center in Cupertino centered on immersive experiences. Unlike last year’s event, which largely centered on real time 3D experiences like *Encounter Dinosaurs*, this one was all about **Immersive**

Video.

**The Apple Vision Pro has a killer app, and it's neglected**

Three months in, I've put my Vision Pro on dozens of faces, and observed carefully. Now I know what will make people...

medium.com

A lot has happened in the year and a half since I wrote the above article: the first commercial camera capturing Apple Immersive Video (AIV), the $33,000 Blackmagic URSA Cine Immersive, finally shipped to dozens of studios. visionOS 26 launched with many new media capabilities. Excitement has been brewing among creators and, despite persistent rumors about the death of the Vision product line, Apple just quietly released a refresh.

I had bet that AIV would be the first thing that really takes off with the Vision Pro and that’s why I started Acute. With this event, Apple signaled its desire to foster an immersive video ecosystem and confirmed their commitment to the Vision line.

The 10 hours of sessions from both days are available on YouTube. Below are my notes, takeaways, and impressions, focused on AIV and other immersive media formats. I hope they’ll be a useful resource to the community.

> TL;DR: Apple and Blackmagic greatly streamlined the immersive video process to make it easier for industry professionals. But the devil is in the details, and, boy, are there many details!

. . .

## Day 1

The auditorium was half-empty when the first session started at 10am; last time it was full. A quick show of hands showed that half of the audience were first timers. I was afraid this was a sign of attrition in our space, but the room filled up to ~3/4 by the end of the morning.

## Tim Dashwood — What’s new in immersive media

This session largely overlapped with the WWDC25 session about video experiences for visionOS, with an updated state of software readiness.

### Per-frame Dynamic Masks

New in visionOS 26, all video formats now get a per-frame dynamic mask designed to eliminate letterboxing (black bars) in video playback. Transitions are made seamless with a resizing screen window.

### Spatial Video

There’s a persistent confusion between *3D video* (rectangular stereoscopic, made for theatrical release), and *Spatial Video* (rectangular stereoscopic, captured with iPhone).

Spatial Video is an Apple format that comes with extra metadata and rendered feather edges to maximize viewing comfort in the Vision Pro.

Traditional 3D video lacks the metadata and therefore must “bake” the viewing comfort, especially around the edges, in the pixels.

Tooling support for Spatial Video is expanding: it’s now supported by Compressor, DaVinci Resolve Studio, and Final Cut Pro.

### Spatial Scenes

Spatial Scenes are a new feature of visionOS 26, they’re mini-dioramas AI-generated from photos. Unlike *Spatial Photos*, they’re not a stereo image with a fixed point of view, you can look around them.

They have some limitations (RealityKit only, they need to be generated at runtime and cannot be persisted, they don’t animate), but some apps like Caradise already integrate them.

### APMP

The Apple Projected Media Profile is Apple’s way of supporting VR180,

VR360 and other exotic projected videos by injecting metadata into them. I covered it here and at the Immersive Video Enthusiasts meetup, .

Tooling support for APMP is still very limited, but macOS 26 Tahoe introduced two new ways to inject the metadata in existing videos:

- The **avconvert** command line built into macOS (or is it built into Xcode's command line tools?) now supports APMP. You will need to set its parameters properly (check `avconvert --help`).
- Finder has an “Encode Selected Video Files” option when right clicking a video with APMP options.

On my own machine (macOS 26.0.1) however, I got an error applying one of the suggested MV-HEVC presets to side-by-side or over-under videos.

### AIV

Apple Immersive Video, covered in details in later sessions, had a couple announcements:

- Live AIV streaming is coming up, with live LA Lakers games coming early

2026. Our friends at SpatialGen spent a bunch of the summer developing support for live AIV.

- The *upcoming* version of Compressor will be able to generate AIV HLS fragments, the chunks of video used for streaming.

### AIV Static Foveation

Apple lifted the veil on **Static Foveation**, an important ingredient to image quality in AIV.

The camera captures 8160x7200 per eye, but frames stream at 4320x4320 per eye. Static foveation is about distorting the original fisheye projection to preserve the pixels per degree (PPD) around the center of the image, at the expense of details at the edges. The parameters of static foveation are encoded into the video’s metadata track for seamless unwarping at playback time.

Apple says tooling is on the way, with support in DaVinci Resolve Studio, Colorfront, SpatialGen, and others coming in the next few months.

It is strongly suspected static foveation was used in all of Apple’s AIV videos,

where it was achieved programmatically, which is part of why those videos still look sharper than Blackmagic URSA Cine Immersive-made videos.

## Adarsh Pavani — What’s new in interactive

This session focused on new features in RealityKit and SwiftUI for visionOS 26. It was interesting but I’ll keep this article focused on video topics.

## Elliot Graves — Creating with immersive video and audio

**If there’s one talk to watch, I recommend this one.** Apple is known for their sleek presentations, and this one is a great example of a wealth of learnings being synthesized, organized and delivered in an effortlessly clear manner.

While not entirely a new discipline, VR/immersive video is still very new and poorly understood compared to traditional media. A lot of trial-and-error is being done on sets, in post-production, but also with audiences. Apple’s productions themselves are imperfect at times, and educational in that capacity.

### Key Foundations for Immersive

Great immersive videos maximize and preserve four aspects: **Presence**, **Authenticity**, **Proximity**, and **Connection**.

### Presence

Presence is giving the illusion to the viewer that they're teleported. It's one dimension of the *suspension of disbelief* that's particular to experiences in the Vision Pro, where visual acuity and audio realism make you forget your eyes are looking at a screen.

To achieve a "high fidelity of presence", the video expands to include peripheral vision, and scale and depth are accurate.

This means that traditional 2D film composition techniques like the rule of thirds, golden ratios, etc. are no longer relevant: what's important needs to be at the center of the frame.

The eyes of the audience are guided differently. With depth, the composition has to contend with foreground, middle ground, background. Height should be a reasonable standing or sitting position. The horizon should be level. Leading lines and concentric circles are effective guides while preserving the viewer's sense of agency.

### Authenticity

Authenticity starts with depth, scale, and acuity matching the viewer's expectations, but it shines with *creative motivation*.

With such high resolution and the viewer's freedom to look everywhere, nothing escapes the audience, from subtle body language on the talent to the artifices employed to serve the narrative. Editing cannot masks flaws as effectively as it does in 2D, so subject and performances should be picked very carefully.

Preserving the viewer's agency to look anywhere is to preserve the authenticity. They can still, however, be guided with audio.

### Proximity

Proximity is about feeling physically close to the story. With a strong sense of presence, audiences will remember an immersive story as a personal experience rather than a film they saw.

The camera cannot zoom, it has to be positioned thoughtfully. You can take cues to the 4 P’s:

- Public (8m~25ft) for strangers;
- Professional (4m~12ft);
- Private (2m~6ft) for close friends;
- Personal (1m~3ft), within arm’s reach.

Unlike real life, the audience cannot step back when it’s uncomfortable. So close proximity must be used sparingly and intentionally: keep subjects at 2m~6ft at least. Think about how characters are introduced, and give the audience time to get comfortable with them as you reduce their distance.

### Connection

Connection is about making the audience more than a mere fly on the wall. A powerful tool for connection is eye contact. Another one is motion.

### Egor Ananyev — Motion and Immersive Media

1 in 3 people are subject to some form of motion discomfort in day to day life: it’s called the *visual-vestibular conflict*. It happens when our eyes and internal ear disagree in their perception of motion, that’s what makes people sick in VR.

It’s not only about moving the camera: large objects in motion, or many small object in motion together can trigger motion discomfort. It’s about “motion intensity”, in other words how much of the visual space is in motion at any time.

It’s more comfortable to see motion if the objects are smaller and further. Texture and contrast can affect comfort too: simple texture make motion less perceptible. Predictable motion is also more comfortable, like a smooth, slow pan along a single axis.

Motion can be counterintuitive: slow but close to the ground can be more intense than fast high in the air.

visionOS 26 auto-detects excess motion and, depending on system-wide user settings, can automatically dim or even pause the video to preserve viewer comfort. Immersive video can also be viewed through a rectangular portal for most sensitive viewers.

Those system features exist at the expense of immersion and presence, so the best place to ensure viewer comfort is while filming.

Editing can help too: symptoms of motion discomfort subside over time, two moving shots in a row will feel more comfortable if a still shot is inserted between them to “reset” the viewer.

### Conclusion

Elliot concluded the session noting that motion should not be avoided altogether, but should be employed when it serves the story, with precision. Elevated, one of Apple’s shows, is centered on exploring motion in immersive video.

### Missed connection

An important topic that was skipped over, in my opinion: *who is the viewer?*

As audiences get thrown into a different place and a different time, scripted

productions need to clarify to the viewer who they're supposed to be.

The *Adventure* series does it extremely well: the voice over invites the viewer inside the head of the episode’s hero, like Faith Dickey in the *Highlining* episode. The viewer gradually *makes one with her* and appropriates her struggles and wins, feeling her emotions.

Another successful example is Immersive Flashback’s immersive video of Bastille Day with the French President, which I was lucky to watch at the event.

In the first shot, Emmanuel Macron sits across from you and welcomes you to take part in the celebration, identifying your role as some sort of VIP guest. Everything from there makes sense, the bumpy ride on the military jeep or your seat among dignitaries in the presidential stand.

I'll say this was the weakness of Edward Berger’s *Submerged*. The viewer is never identified or acknowledged, and it erodes the suspension of disbelief to a point the whole video feels like a behind-the-scenes montage of a Hollywood shoot. A very good one, for what it’s worth, but this wasn’t the

director’s intent.

## Serenity Caldwell — Design Immersive & Interactive Experiences

This session presented good practices that are applicable to any Vision Pro app.

### Familiarity and pacing

A core message was that this space is so novel that any apps, and by extension any immersive video, *might be someone’s very first experience with virtual reality*. Therefore, any app should start from a place of familiarity.

A good example is gradual immersion: start with a familiar window before throwing the user into a fully immersive experience. Another is pacing: build trust with the audience, guide them, set expectations before smoothly transitioning to big moments.

Video apps should should follow the analogy: they should open to a “flat” catalogue window reminiscent of TV apps, but once the Play button is pressed, the immersion can wrap all around the viewer.

An important idea pointed out is that *not all interactions are active*: passive/limited interactions like looking around, staring, listening, entering or exiting a mode *are* interactions by their nature of being intentional movements of the hands or the eyes.

### Victor Agulhon — Insights and Learnings from D-Day: the Camera Soldier

Our friend Victor from TARGO, whose decade-long work in VR video was a personal inspiration for me to get into this space, recently launched an ambitious (and amphibious) app, D-Day: the Camera Soldier.

It mixes moments of archive footage, present time immersive video, 3D immersive environments and interactive 3D objects.

A key takeaway was that through early access testing, they learned that those modalities had to be used with purpose and restraint. In particular, *attention is more important than actions*.

### Audio in Immersive Environments

Serenity took back the stage to talk about the sound design of immersive environments. This certainly applies to video as well.

A great recipe is to layer multiple types, as done in the Mt Hood immersive environment:

- looping ambient sound (e.g. wind, rain) in the background;
- looping spatial sounds (e.g. chirps, foliage), muted and at a distance, something that could come from ambisonic recording;

- nearer and crisp spatial sounds with randomized intervals and locations so they don't feel repetitive (e.g. croaks).

Audio can set expectations: at the right volume and distance, danger can be exciting while still comfortable (like an alien lurking in ventilation shafts).

Score and sound effects can set a tone and direct attention, like how the whimsy music at the beginning of *Encounter Dinosaurs* fades to a distant thunder, causing the user to focus on the portal that just opened, and setting expectations for elevated danger.

## Nathaniel Ellingson — Build spatial experiences with RealityKit

This session went over some of the more recent features for 3D spatial game development. One of the highlights was how to set up a portal effect with 3D content poking out of it, which could be a neat trick for mixed video/3D content.

## Alex Rodriguez — Build shared experiences

This session was all about SharePlay, Apple’s system built on top of iMessage for synchronizing applications across users and devices.

SharePlay is deeply integrated into Apple’s other frameworks like AVFoundation, the API for video playback, which means that it’s notionally easy to synchronize videos across multiple Vision Pros, with built-in support for play, pause, seek, etc.

What exactly is synchronized vs. managed locally is up to the developer. The key classes to use for SharePlay are `GroupActivity` and `GroupSession`. Once a session is created, `AVPlaybackCoordinator.coordinateWithSession(_:)` ensures that playback is sync’d across participants.

Our friends at Vantage extended OpenImmersive to allow SharePlay-based boxing watch parties.

Interestingly, the new default player for AIV hides the virtual personas. I think it’s a shame, I felt more invested in the social dimension of the watch party seeing the other guys’ avatars.

## Q&A

The presenters got brought back on stage to answer audience questions.

### How to publish AIV to other platforms that don’t support it? (Samsung, Meta, etc.)

DaVinci Resolve Studio has an export preset called “VR180” that exports AIV as a side-by-side half-equirectangular video, which is compatible with all VR video platforms. AIV is the go-to format for the highest quality however.

### What’s the difference between stereo VR180 and AIV?

It’s in the metadata. VR180 is half-equirectangular video and that’s it. AIV maintains the fisheye of the original image, embeds a lens calibration file (known as .ilpd). The image is undistorted and carries the metadata all the way from capture to playback.

I’ll note that the upcoming *static foveation* contradicts the notion that the image is undistorted all the way to playback.

### Where to find samples for AIV and other APMP formats?

HLS sample streams are available on the Apple developer website; BRAW sample footage is available on the Blackmagic Design website.

### How does SharePlay for AIV differ from other experiences?

Since AVPlayer supports AIV out of the box, AIV is supported by SharePlay like any other audio/video media, with join, catch up, seek, etc. There are UI indications when participants join and leave. The Spatial Personas are hidden, but you can still hear participants talk.

. . .

## Day 2

Last year’s event was a single day; this year’s allotted a full day for Apple Immersive Video specifically. The crowd size was about the same as Day 1, showing that like in previous events, participants were primarily interested in AIV.

## Ryan Sheridan — Apple Immersive Video: a fundamental shift for video and audio

This dense session went into the weeds of *what is Apple Immersive Video?*

The goal of AIV is to mimic 20/20 human vision, for 180–230 degrees of field of view (FOV).

### Capture Specifications

AIV uses *“dynamic bespoke projection”*, meaning there’s no default projection,

no conversions to and from equirectangular latitude/longitude. Instead, lens metadata is carried all the way to playback.

AIV is captured in world-scale (metric) accuracy, there’s no warping or stitching. There’s an *Immersive Lens Profile Data* (.ilpd) calibration file per shot, which is a small JSON file of 50KB or less, containing everything needed to reproject the frames. There’s also a lens profile file generated during factory calibration that’s attached to all clips.

The AIV format mandates minimum capture specifications that are very close to the Blackmagic URSA Cine Immersive effective specs: 90 frames per second, stereo with 7200x7200 pixels per eye. This is 44 times the data of a typical 2D cinematic capture.

Traditionally video is expressed in “K” values: 4K, 8K, 16K, etc. AIV is expressed in acuity, striving to be as close as possible to 20/20 vision, which is about equivalent to 60 pixels per degree (PPD). Capturing 60 PPD at 180 degrees would require a sensor size of 10800x10800 pixels per eye. At 7200 pixels, we are at 40 PPD, incidentally close to the peak acuity measured on the Vision Pro display.

### The dual AIV format

The AIV workflow strives to streamline the complexities that veteran VR filmmakers have had to contend with for over a decade.

The all-in-one camera delivers a single file that’s kept through the whole post-production flow all the way to encoding and delivery.

AIV is a “compound format” that assembles a number of elements under the hood: all AIV files build in video, audio, and other metadata.

But there’s a subtility: **AIV comes in two flavors, one for production, and one for delivery (distribution).**

In my opinion, this duality will be a big source of confusion for the years to come. Apple strongly assumed that creators are industry professionals who hand over their (production) AIV to a distributor who then encodes it into a (delivery) AIV.

In practice, production AIVs are too big and impractical to preview, share with collaborators, etc. Furthermore, it seems like a majority of AIV creators are smaller studios and indies, and they tend to wear all the hats themselves. Lastly, AIV is a format where visual fidelity is the main selling point; studios will want to preview the quality delivered to their audience early in the process.

**Production AIV files** are huge (terabytes!), mostly because they package video data in RAW (such as Blackmagic .braw) or intermediary formats (such as QuickTime ProRes .mov). They can also include multipart VFX in .exr or full on .usdz 3D models to use as immersive environments, to replace the black edges of the screen.

**Delivery AIV files** (or .aivu) are *merely* big (tens of gigabytes): they compress video into MV-HEVC, an efficient way to pack stereo imagery. Similarly, their audio and metadata tracks are compressed from the raw sources in the production AIV.

Production AIV with raw video, audio and metadata to the left, encoded into delivery AIV to the right.

### Best practices for AIV delivery

Preserving visual acuity is the most sensitive part of producing a delivery AIV. In Apple TV, AIV is delivered over HLS (streaming), at a resolution of 4320x4320 per eye. Apple discourages going over this resolution, because the hardware won’t decode it in HDR at 90 frames per second consistently. Likewise, compression for streaming needs to take the bitrate down from multiple Gbps to under 100Mbps. Efficiency is the name of the game, to make the most out of the pixels and the bytes.

Video compression isn’t “smart”, it tries to preserve details evenly across the image, with a bitrate target. This means that loss of detail will happen evenly, which is why **noise is the enemy**. On top of being a source of left/right discrepancy, from a compression standpoint, noise is a legitimate part of the image and the encoder will spend its “bitrate budget” trying to preserve it. At 90 frames per second, especially at higher ISOs, noise is everywhere.

**At Acute, that’s our number 1 piece of advice: denoise aggressively.**

Denoising can give the image an overly soft, plasticky feel. It’s a good idea to pair it with **edge sharpening and detailing**, for which more and more AI-based solutions exist. All this is computationally costly but will make encoding preserve the acuity for your audience.

Another aspect, mentioned earlier, is **static foveation**. Linearly downscaling the frame to 4320x4320 results in 24PPD, so warping the frame to preserve 40PPD in areas of interest (typically, the center) while degrading the acuity at the edges makes the best use of the pixels. Static foveation is not available in today’s tooling, but it will soon be possible to apply it per-clip.

The last aspect is the **encoder**. Right now, DaVinci Resolve Studio exports using the *hardware encoder*, that is to say the fast but inefficient part of the GPU chip that’s hard-wired for video operations. *Software encoders*, whether they run on the GPU (like Apple’s VideoToolbox), or the CPU (like x265), are much slower but result in much smaller files.

Acute uses SpatialGen’s software encoder, which is efficient, and for this reason we ask our creators to upload exports from DaVinci Resolve set at a bitrate of 200–300Mbps.

### Deep Sen — Meet Apple Spatial Audio Format and APAC

**Apple Spatial Audio Format (ASAF)** is the “raw” representation of audio within AIV.

It’s more than just an alternative to other “surround” or object-based audio formats like Dolby Atmos, it combines different types of audio channels: **high order ambisonics** (HOA, which you can think of as the audio equivalent to 360 mono video), **objects** (each emitting sound with a position and an orientation relative to the listener), and regular channels (e.g. stereo).

A bit part of ASAF is its audio renderer, which mixes all these sources in real

time to play them accurately on the listener's hardware. In the case of the Apple Vision Pro, those are the two stereo speakers, whose output depend on head orientation (known as *binaural*).

**Apple Positional Audio Codec (APAC)** is the compressed representation of audio for delivery. A typical setup of 5th order ambisonics and 15 individual objects compress into 1Mbps / 768 Kbps / 512 Kbps for excellent / good / fair fidelity, representing effectively ~1% of the overall bandwidth of AIV.

APAC workflow tools are coming out: the main two at the moment are the built-in support in DaVinci Resolve, and the new ASAF suite in Apple's AAX Plugins for ProTools. Both come with a 3D panner for positioning objects in space, and a video player to overlay the objects on the frame.

. . .

## Production workflows for Apple Immersive Video

These sessions gave a bird's eye view of the tools to capture and edit AIV.

### Austin Novy — Capturing Apple Immersive Video and working on location

Austin shared advice from his experience shooting on the field with the Blackmagic URSA Cine Immersive, on unscripted shoots in difficult

environments.

- The camera is big, so when going outdoors you want to keep the footprint small. Use <100W travel-safe batteries in stackable V-mounts for easy in-and-out of the backpack.
- For audio, a multitrack recorder like the Zoom F6 is an essential accessory to get all sources on the same time code and save time in post.

- Allow for enough media for the creative vision, especially in unscripted where you don’t want to miss a great moment because the camera wasn’t rolling: there are up to 16TB modules (but beware: big modules = big downloads of dailies). Overpack so that battery and storage are the last thing on your mind.
- For monitoring, the camera's LCD is good when you need to be nimble and get rolling fast, but it can be set up with external monitors. Most crew will work with a single eye view, but dual eye is important to check for consistency, flaring, cropping.
- Take time to prepare the camera, fix the horizon. The tripod is the most used tool: while static, it still lets the audience creatively explore the field of view. The camera is also compatible and can be mounted on other things: FPV drones, scorpio cranes, even steadicams. Helicopter and

underwater housings tend to be custom.

- Check the work with dailies. It can be done on the go on a MacBook Pro, but if live color or fast backups are needed you need a cart with multiple Mac Studios and multiple monitors.

I’ll add here that several creators I’ve talked to have invested $30–40,000 of hardware and accessories on top of their $33,000 Blackmagic camera.

### Alex Weiss — Spatial Audio workflow for Apple Immersive Video

Alex covered the three main phases: **acquisition**, **sound editorial**, and **spatial mix**.

- During acquisition, it pays off to record ambisonic from the camera’s perspective. It’s the foundation of the video’s soundscape. Even if you use only one channel from it, you have options in post.
- Mono recording still matters: record as many mono sources as possible, one for each talent. You can’t use booms because they get in the frame, so prioritize individual lavalier microphones.
- During sound editorial, as mentioned before, the two main tooling options are ProTools and DaVinci Resolve’s Fairlight feature, where spatial audio is deeply integrated. Assemble all productions sounds and start layering.
- Immersion is in the details: layer Foley and ambiances on top of production capture.

- Spatial mix is where all the sources are combined and spatialized. Use visual anchors to map the sounds to visual cues. It’s critical to mix binaurally (that is to say, with headphones that support spatial audio), and review immersively in the Vision Pro.

### Dave Hoffman (Blackmagic Design) — Apple Immersive Post Production Workflows

Dave gave an overview of the immersive features of DaVinci Resolve.

- The camera’s media module can transfer files through WiFi, Bluetooth, Ethernet, but also slot into a dock and mount as SMB.
- Resolve has a 2D “immersive viewer” where you can drag the mouse to look around; it also offers live preview on Vision Pro.

- You can also visualize the curves of motion data (accelerometer and gyroscope), which is useful to trim those parts out for motion comfort, or sort the clips as motion can be an emotional beat to the story.
- The timeline shows a new type of track, the “background track”, which is rendered beyond the boundaries of the video track. By default it’s black, but it can load a .usdz environment. We haven’t really seen this in the wild; I wonder if that’s how Bono: Stories of Surrender mixed Spatial Video and 3D graphics.

A few other important tools unique to AIV workflow were highlighted:

- **Edge Blend**, which is a fine control of the frame’s feathering. It can be used creatively to focus audience attention to a narrower field of view, or pragmatically to hide equipment, or a tripod leg.
- **Immersive Patcher tool** (in Fusion): it temporarily undistorts the image, which is useful for painting over the frame or running tools that assume rectilinar images, like stabilization.
- **PanoMap tool** (in Fusion): similar to the Immersive Patcher, it converts the image to and from various representations, like equirectangular (lat/lon).

- **IntelliTrack Panner Tracking**, a tool that uses AI to track objects frame-by-frame. It’s compatible with the sound editing pane Fairlight, for easy mapping of spatial sounds to moving objects.

Lastly, the delivery page has multiple presets, which have been source of confusion:

- **Vision Pro Review** generates a **delivery AIV** that’s a smaller file meant to be airdropped or transferred to Vision Pro using the Apple Immersive Video Utility. The reason it’s called “Review” and not “Delivery” is because Resolve uses the hardware encoder, and therefore the generated file is not particularly optimized.
  However, this is the right option if you’re uploading it to Acute for distribution, you’ll have to make sure to set the bitrate high (200~300Mbps) so that details are not crushed by compression.
- **Vision Pro Bundle** generates a **production AIV** that’s a large file where the video is encoded in ProRes for archival or further encoding via Compressor or other tools.

A future preset with efficient encoding for delivery was teased, stay tuned for that.

### Brandon Heaslip (ColorFront) — Dailies for Apple Immersive Video

ColorFront is a company that produces tools for the movie industry. They’re introducing AIV workflows support.

- **On-Set Dailies** is a suite for color-accurate review.
- **Trankoder** is a mastering program for delivery. It supports exports to AIVU for visualization and ProRes for editing, distributed rendering across multiple machines from a render queue.

In their next update, they’ll integrate the MainConcept encoding library to

the platform, which will greatly increase compression quality.

### Zach Handshoe (SpatialGen) — Self-distribution Apple Immersive Video

SpatialGen is an encoding and streaming platform for immersive video formats, used by many Vision Pro apps, and also by other platforms such as Acute.

They recently introduced a desktop uploader to manage large uploads such as the AIV ProRes Bundle out of DaVinci Resolve. In their next update, they’ll support static foveation.

I can’t say enough good things about our friends at SpatialGen. Zach describe their service as the last mile of distribution; I’ll plug that Acute is the last mile of the last mile, as a CMS for immersive video and app template.

## Lessons from Apple Immersive productions

These sessions brought together several of the people behind the AIV releases from the past year, sharing learnings for a wide range of production types: music video, scripted fiction, live events, unscripted documentary.

### Anton Tammi (The Weeknd: Open Hearts) — Benefits of pre-visualization

Anton directed the first ever AIV music video. He had directed most of The Weeknd’s hits, but having no prior experience with VR or immersive, he had to come up with creative ways to make his production a success.

I’m personally a fan of this video, because it manages to be highly stylized, with heavy color grading and a lot of VFX without hurting the sense of presence, and it uses a lot of motion without generating any discomfort.

The key was **pre-visualization**, a “modern storyboard” made of low-fidelity CG animation to experiment and compose the eventual shots.

His team created custom camera rigs in a video game editor to mimic the immersive camera’s lens characteristics, and visualized the renders in the headset before shooting anything.

He found a number of things that way:

- 2m~6ft is a good minimum distance to avoid claustrophobia or discomfort in shots: a sports car was too cramped, the back of an ambulance was comfortable.
- Rotating pans create motion sickness, but plenty of motion is still possible: slow pans in a corridor, fast drives with a car hood in view, environment distortion in the distance.

Pre-visualization helped him greenlight his ideas with the talent, and educate the team on what was expected in production: set design, light setup, camera movement, etc.

His takeaway: immersive pre-production takes more time, but it’s cost effective once you factor the communication and preparedness benefits in production and post-production.

### Alex Miller (Submerged) — New technology and production expectations

Alex was a producer and presented her takeaways from working on the first scripted AIV fiction, Submerged.

She highlighted some of the differences compared to 2D filmmaking, emphasizing the earlier stages: **creative development** and **pre-production.** They got the crew up to speed, making sure everyone was familiar with navigating the Apple Vision Pro.

Like Anton, they made extensive use of pre-visualization and realized that the real dimensions of a WWII-era submarine would make the experience uncomfortably claustrophobic, so they pushed the walls a bit.



They also found that the very high resolution of capture meant they couldn’t cut corners on decors and costumes: fake props, fake tattoos were immediately noticeable: they had to change all the dials from plastic to period-accurate brass and glass to preserve the immersion.

Frequent on-set reviews in the Vision Pro was key to catching these problems before they would be sent to post. They even had a crew member on-set whose full-time responsibility was ensuring the devices were clean and charged at all times. Whew!

*Submerged* is often brought up in conversations in the immersive video community, because a lot of people feel like something didn’t quite work, but it’s hard to put a finger on it.

My guess is that authenticity is hard to achieve in AIV fiction. Maybe it’s like the motion smoothing on your parents’ TV, we’re so used to reality looking a certain way, and cinema looking, feeling another way, that anything in between feels uncanny and inauthentic.

### Ivan Serrano (Metallica) — Lessons: access and capturing live events

Ivan is a technical director in Apple’s immersive media team. He shared takeaways from shooting Metallica’s concert at a stadium in Mexico City.

Traditional live event captures rely on cameras with long telelenses shooting from afar, and a couple shoulder cameras, to stay out of the way of the performers and the audience. Immersive cameras cannot zoom, cannot move much either. In order to create proximity, the challenge is *access*.

Access, in order to set the cameras as close to the action as possible, starts with negotiations: security, broadcasters, etc. Like in the previous examples, creative development groundwork is essential. The team traveled to meet the band, share concepts and demos, and get their buy-in.

Most immersive cameras would be static for the whole show, so positioning and orienting them well was critical to gather great footage. The team studied the stage layout (a wide ring with a crowded “snake pit” of fans in the middle) and how the various band members moved around it across multiple shows of the tour. From there they identified the spots where the big moments would happen: solos, crowd surges, moments of intimacy. Like The Weeknd, they used 3D pre-visualization to stage the shots.

In the end 14 cameras were deployed around the stage, including an overhead spidercam. Backstage, the event was live directed like a broadcast even though the video would be later post-produced and distributed on-demand.

*Metallica* is one of Paul’s favorite immersive videos, and that’s proof a well executed production has appeal beyond a fan base. You absolutely feel the energy of the crowd (and some of the vibrations of the drums).

### France Costrel — Unscripted productions and uncontrolled environments

France is a creative director at Apple, and her team produced many of the AIV documentaries, such as the *Wild Life*, *Boundless* and *Elevated* series. Those documentaries are shot by smaller crews, in largely uncontrolled outdoors environments.

She emphasized simplicity in designing the shots, because the viewer is ultimately in control of tilting and panning. However, simple doesn’t mean easy, and her team found rule of thumbs for composing shots:

- 2m~6ft is a comfortable camera distance to its subject, you can go down to 1.2m~4ft if you want to feel close.
- 130cm~50in to 150cm~60in is a good range for “collar bone” camera height. Getting height right really matters.

Rules can be broken: in *Wild Life*, a rhino gets way too close to the camera. The shot is a fan favorite, but that’s the only shot that does that in the whole video.

The environment doesn’t always give you a choice. In *Boundless*, they had to put the camera in a hot air balloon and the structure of the basket felt too close. They kept those shots brief and focused the viewer away from it by focusing on the pilot.

Immersive video doesn’t have its own cinematic language yet, but some of her findings already echo those presented by Elliott the day before:

- Motion helps the sense of presence if done mindfully.
- The 2D rule of thirds doesn’t apply, in immersive you compose with foreground, middle ground, background.
- Viewers can be guided and directed with spatial audio.
- Casting of the talent is fundamental. You need to establish a relationship of trust, because any anxiety will translate into body language, little twitches that are immediately obvious to the viewer. Overconfidence is not good either, since the pace is slower, viewers spend more time with characters, they need to be likeable.

My favorite takeaway: sometimes, voice is not needed. By giving the viewer

the time to look and experience, they get to connect and share the moment in their own terms.

### Tim Amick — Lessons in presence in post production

Tim is the head of post-production in the Apple immersive media team. A common theme across the event was reassuring the filmmaker audience that *the process isn’t that complex* and in fact mostly similar to 2D editing.

The main difference he sees is that it’s important to **preserve the sense of presence**, echoing Ryan’s earlier presentation about the AIV format and workflows.

His presentation was rich with insightful tips:

- Check shots for lens flare; but if the flare is on both eyes it can be OK.
- Avoid overloading the frame with information, things to look at: the viewer cannot look at everything at once, and may feel they’re missing out on a key story piece.
- Create the illusion of freedom for the viewer to look anywhere, but still direct their attention.
- Consider comfort in eye convergence when cutting shots, e.g. from wide

open to a closeup. A quick fade to black can help reset the viewer to the center.

- Mix spatial audio carefully: in real life, our brain filters out audio cues; in immersive, it’s the audio mixer’s job. Sound reinforces what you see. Ambisonics can take you most of the way there, but Foley and sound design are still needed to match our subconscious expectations.
- Audio position is very important, e.g. voice needs to match the position of the mouth. If missing, even non-technical people will tell something is off.
- Similarly, CG will be under heavy scrutiny, it has to look very real. Flaws will stick out immediately to all viewers.
- Every shot needs denoising, for the sense of depth and to help with compression. Compression artifacts will **really** stand out in immersive.
- When color grading, an idealized reality is a strong starting point. Start realistic and push to the limit where it still remains authentic. Beyond that you can break the illusion. Validate color grading inside the Vision Pro.
- Everyone who makes creative decisions must review in the Vision Pro, so many things can be missed on a 2D screen.

## Q&A

### How did Apple decide to embrace 180 rather than 360 degree video for immersive?

Apple developed both AIV and also APMP to cover other immersive formats. For AIV, 180 brings two big strengths: the audience gets less FOMO and doesn’t need to twist as much; and production is easier, it’s hard enough to hide crew and equipment in 180.

### What tools support ASAF?

The two options at the moment are Apple’s AAX plugin in ProTools, or the

integration in DaVinci Resolve's Fairlight.

### How to get started with an ASAF project?

The ProTools plugin has a template and some tutorial videos.

### Is there a way to create custom .ilpd files for 3D VFX workflows?

They're created at the factory during calibration, but some exist for fixed projections, such as equidistant (fisheye). Those can be used with VFX rendered with the same projection.

### Are there best practices for getting started with CG content?

A lot of same principles of pre-visualization for capture applies: whether you do full CG or integrate it with capture, it's always best to start by mimicking the Blackmagic camera characteristics (60mm IPD, 180 degrees). Experiment, review and iterate as much as possible.

### What encoders take advantage of static foveation?

SpatialGen announced that support is on the way and we expect more developers to announce upcoming support in the coming weeks.

### How to go about shooting immersive interviews, for instance with multiple people? Should you use multiple cameras?

Shoot with only one camera, since it sees 180 degrees it will capture everything. But be mindful of not making the viewer feel left out: work out the right distance, and let the speakers occasionally glance at the camera, eye contact engages the viewer. But be mindful that stares can be awkward and make the viewer feel that something is expected of them.

Since there's only one camera and the conversation flows organically, there's also a lot less cutting and editing. It's important to brief the characters that everything they say, however they say it, is what the video is going to be, and they should be leading this.

### How to bring non-Apple immersive media into AIV projects, like 2D archives, or other camera systems?

It comes up a lot. In DaVinci Resolve, you have two main ways, either use the Fusion tools to render the footage at their highest possible resolution and fidelity, or you can integrate it in a VFX pipleine, for instance project it on a curved screen, add some lighting on a fake floor, etc. Ideally, use it sparingly because your audience may wonder why they're getting in and out of immersion.

### Is it possible to define the position of static foveation shot to shot?

This is a feature of encoding, so it depends on what tooling the developers offer. At first it won't be "areas of interest" quite yet (it will be a fixed forward function) but we'll see what developers come up with.

What I think Apple is trying to say is that yes, the *AIV format* does allow shot-to-shot foveation changes, but the tooling to achieve them is going to start rudimentary with presets. I'm pretty certain that not only static foveation was used in Apple's productions, but it was achieved with custom software engineering, hence Ryan's answer essentially saying "it depends what tools developers end up making".

### Is there something you'd like to try or see with AIV?

**France**: I'd love to try using a steadicam to follow a character around and see the world like they see it.
**Alex**: Two things: One, I want to experiment with spatial audio storytelling, eliminating or reducing voice overs and score, and playing with dynamic range. Second: there's so much to try in music, like with *Alicia Keys: Rehearsal Room*, blur the lines between concert and being on stage intimately with the artist.
**Ryan**: I want to go sit in the places where I'll never get to in my life (the Moon, the Titanic's deck, etc.), experience them, just sit and enjoy the space.
**Sarah**: there's a lot more world than where we've already been to with the various shows, more stories waiting to be told in new places. I want more!
**Matt**: it's such a great format for connecting with people, I'd like to see that done over a long period of time, which means the production would have to

run over years. Also I’d like to see a post-production stick to schedule once!

. . .

## Workshop: Post production for Apple Immersive Video

Two hours of deep dive tutorials closed the day, covering editing, editorial, visual effects, spatial audio, color grading. Some of the examples echoed the previous sessions and put them into practice.

There was no video recording of this, so take my notes with a grain of salt!

As a secondary source, consider the latest workflow guide for immersive from the Blackmagic support page. Our friend Hugh Hou is also in the process of putting out a series of YouTube tutorials. For more general guidance, look for Mike Swanson’s long form Immersive Video production tips.

## Justin Beach — Editing for Apple Immersive Video

Many tools and techniques in traditional editing are still valid, but you have to cut and pace a bit differently.

### Consider gaze and curiosity

Find where audiences are likely to look, it is critical for smooth cuts. The position of gaze carries over from a shot to another, be careful not to lead gaze to an empty spot: viewers should not spend time searching for where the action is.

About curiosity: exploring the scene takes time, take it into account in pacing the shots. The audience will feel cheated if they don’t have time to explore.

### Consider depth and motion

Take into account that your audience's eyes have to converge to subjects. Be careful if you're cutting between far and near shots, it could be uncomfortable. A possible technique is to insert intermediate depth shots to make it gradual.

Beware of compounding motion. A static shot between two moving shots will help the audience reset.

## Clara Chong & Ben Allan (Main Course Films) — AIV Post-production seminar: Editorial

A lot of same principles from film, VR, and real life apply.

How viewers feel the storytelling come from the "three P's": *proximity*, *positioning*, and *perceptual salience* (what draws attention). For those, extra planning is key. AIV is a new creative language, and needs new creative thinking: choose subjects and locations that work for immersive.

Main Course Films shot two projects with the Blackmagic URSA Cine Immersive camera and DaVinci Resolve workflow: *The Fine Dining Bakery*, a documentary; and *The Dobos Connection*, a scripted drama.

### Proxy and BRAW

In Resolve, Apple Immersive timelines have two resolution presets: full and half. Half resolution is a convenient, easy to revert option to reduce render and processing times when editing.

If you have enough processing power, the whole workflow works on BRAW directly, but proxys make it possible to edit on a MacBook Pro without external storage. Generating them is very quick and their functioni is exactly the same as the originals.

### Edit in 2D, review in AVP

Do regular .aivu exports to preview in the Vision Pro. You can stream directly

from Resolve, but it’s more effective to work primarily on a 2D screen and check regularly in the AVP.

In the export tab, use the **Vision Pro Review** preset and AirDrop the generated file, it will play natively.

You will use a lot more work-in-progress renders than you would in any format. It’s important to keep them organized and named, with incrementing version numbers.

It’s harder to make judgement calls in the edit; it’s essential to get feedback from someone who hasn’t seen the timeline in 2D, because they would be primed to know where to look.

The bigger the change of perspective, the longer you have to let viewers adapt to the new perspective. You may not realize how much it changes when reviewing in 2D. You’ll sometimes have to swap the order of shots to eliminate spatial confusion.

#### Quirks of immersive

Create space for the viewer inside the scene, it may feel cramped if walls, furniture (and subjects!) are too close.

It’s not obvious what to do with spatial audio when cutting shots mid-dialogue. Logically the audio should come from the right place, but feels wrong if it changes mid-dialogue. The cognitive dissonance is eliminated if the audio is kept at the previous spot.

Clara and Ben have a book coming soon on Apple Books: “Cinematic Immersive for Professionals”

## Tim Amick & Matt DeJohn (Blackmagic Design) — Visual effects for Apple Immersive Video

Spatially aware details stand out in immersive: inserted titles, camera tracking, painted patch up fixes are important to get right, and the work has to be vetted in headset regularly as discrepancies are very hard to see in 2D.

### The Fusion page

**Fusion** is the page for visual effects in DaVinci Resolve. In immersive, the nodes are duplicated: media1 in/out are for the left eye, and media2 in/out for the right eye. You can drag those nodes or use number keys to visualize them in the monitor. The monitors can have a view LUT for reviewing, e.g. from camera log to rec709.

The graph is where to add tool nodes with a right click. Under the VR menu, there are a bunch for immersive workflows:

- **Immersive patcher**: it undistorts the image from fish eye to rectilinear, making compositing and painting easier. You can rotate and adjust the field of view angle (usually 90). Its secondary input is to apply an ILPD.
- **Panomap tool**: it converts between various formats. For instance, immersive to equirectangular (lat/lon) can be useful for workflows like image stabilization; lat/lon to immersive can be useful to compose CG, which is often rendered in equirectangular.

### Workflow example: remove a tripod leg

As long as the tripod is standing on a flat surface, you can paint over it with the same patch for both eyes.

- Start with media1 (left eye): work on one eye first and then bring over the work to the second eye.
- Add an immersive patcher: *undistort* to get a rectilinear projection.
- Changed the angle to look down to look at the tripod.
- Add a paint node, add a few strokes to remove the leg.
- Map that back to immersive space: change the angle to look back up, and add an immersive patcher to *distort*, instead of undistort.
- Add a polygon for the area to replace, and now you have one eye.
- Transfer to the other eye: overlay with a merge node.
- You’ll see an offset. Add a transform before the node to align the ground planes, fine tune in the inspector. As long as the area is small the technique works.
- Take the left paint work and paste to the right with the transform node. The patch is going to look consistent. Similarly, after the paint node, invert the transform and then re-distort.

### Workflow example: remove the lens at the edge

The easy way is to use the edge map to cut out the lens, but you can also “borrow” the image from the other lens to paint over.

- Start with the left eye.
- Use a simple transform to grab the mask image from the right eye.
- Set merge to multiply by mask.
- Repeat the same process for the right eye.

This technique will work if the background for the paint over is flat or distant.

### Workflow example: paint out a tripod leg on an uneven surface

You cannot paint over because of different or varying depths.

- Do a simple paint for the left eye.
- For the right eye, use a **disparity generator.** The tool calculates the difference between the eyes on a per pixel basis.
- Add a clean plate work to create a disparity, and feed into a tool called **new eye**. It lets you remap pixels from eye to eye.

Only the right eye needs to be set up with that interpolation.

This technique can also be used to paint over the lenses when the background is near and uneven, like foliage.

### Workflow example: 2D video composited into the shot

- Separate layers in edit page with the two clips: (2D + immersive).
- Use **immersive patcher** with both inputs.
- Crop to make the canvas of the 2D video into a square, then resize for full resolution, and feed to immersive patcher.
- Connect the immersive input node to the immersive patcher node’s metadata input. Without that, the immersive patcher would use a default mapping, and we want to match the camera’s lens mapping.
- Add a merge node to view the effect in context.
- In the immersive patcher, adjust the position/orientation, and check the result in the immersive viewer.
- Since it is a 2D video, dial it in: adjust the convergence, look in anaglyph mode to see depth (none by default).
- Shift the 2D forward to dial in the depth.

The same technique can be used with titles, graphics, etc.

### Workflow example: VFX pull

- In the Color page, add a color space transform tool. Set the input to match the camera space (Blackmagic Wide Gamut, gamma gen5 Blackmagic linear) to linearize the footage.
- Remove the edge blend.
- In the Delivery page: select export to **.exr**. The key is that the .ilpd is passed into the .exr for the VFX shop.
  Fusion is a page in Resolve, but it also exists as a standalone app used by a lot of VFX shops.
- Create a duplicate image to get a 360 mapping, to realistically lit up the model inserted into the scene.
- Insert the .usdz rendered out.
- Then add an immersive patcher to distort back to the right camera space.

### Workflow example: stabilize camera shake

- In the Fusion page, resize the image down to 1K. Tracking doesn’t need the full image, it will process faster.
- Use the Panomap tool to convert to lat/lon space, because the spherical stabilizer is expecting 180/360.
- Increase the stabilization strength to 1.0 (max).
- It will process fast thanks to the smaller image.
- Then remove the resize for the final result.

- Trim out some seconds if they shake too much.

### Workflow example: custom titles

In the Fusion page, template titles will map automatically to lens space.

For custom titles, set up a graph rig with spherical cameras and the panomap tool.

## Alex Weiss, Dave Lebolt (DaVinci Systems) — Workshop: Spatial audio for Apple Immersive Video

### Sound map setup

Lack of spatial audio breaks the immersion. Design audio for the entire frame, and expand beyond just the visible (also behind you).

In spatial audio, the workflow is often in equirectangular. Put the 180 degree square frame in the middle of a 2:1 rectangle and super impose a grid of 90 degree squares. There, you can easily place objects anywhere around the viewer.

For ambisonics, you can use a heat map overlay to confirm where the sounds come from.

### Binaural and headlocking

With spatial audio, as you turn your head, the sound stays anchored to a specific location: that’s **binaural** rendering.

Sometimes, you want some elements to play back as regular stereo, stick to your ears and follow you: that’s **headlocking**. It’s great for voice-overs and score.

### IntelliTrack

Fairlight is the page in Resolve where audio mixing, bus routing etc. takes

place. In Resolve, timeline position is shared across all pages, which makes it easy to go across tools on the same part of a clip.

Fusion (VFX) features integrate with Fairlight, such as IntelliTrack, which uses AI to track an object move across the frame over time. IntelliTrack can be used to move the source of a mono spatial audio.

## Chris Hamilton & Matt DeJohn — Picture finishing for Apple Immersive Video

### Noise reduction, color grading, sharpening

Noise is problematic in immersive: it’s inconsistent between left and right and disrupts depth; it also hurts compression. Take extra care not to soften image too much. Do an individual analysis for each shot, rather than a global approach.

Aim for authenticity when color grading. The Blackmagic URSA Cine Immersive has less chromatic aberration but some shots will still need some correction.

Add sharpening to accentuate the fine details, but that step should be *after* creative color grading, to ensure the sharpening doesn’t degrade the colors.

### Color grading project setup

In the timeline settings:

- Enable the Apple Immersive workflow
- Set the resolution to **full** (typically).
- Set Color Science to **DaVinci YRGB**.
- Set Timeline Color Space to **DaVinci Wide Gamut**
- Set Output Colorspace to **P3-D65**, mastering for 1000 nits.

The Blackmagic Design support page has a workflow guide for immersive with all the details.

### Color node setup

In the Color page of Resolve, the 3D panel will be active because the Immersive workflow is enabled; headset preview is also available.

- Insert a first color space transform node with the right input/output.
- Insert a second node for color grading.
- Insert a third node with another color space transform to match your monitor.
- The AVP display is 108 nits, so you’ll need to tone down. Insert a fourth nodes with a color space transform node that only sets the max output nits to 250 (no changes to input/output color space/gamma).

- Judge if details are lost. Adjust the color correction if needed. The final step will be to tone map from 250 to 108 and add an adaptation offset.

If you’re color correcting on the AVP directly, forgo the above setup and map directly to 108 nits.

### Edge blend and World pose

Be careful about stereo discrepancy between left and right feathering. **Edge blend** is applied per-clip, but you can export the feather nodes as JSON to reuse in other clips.

In the 3D panel, use **World pose** to adjust the orientation of the image in lens space, to adjust a bit of tilt and roll. You can also use World pose to flip the image horizontally, which is a useful trick to maintain viewer gaze from shot to shot.

With those effects, tick “visionOS effect” to encode them as metadata instead of baking them into the pixels, which can soften the image.

### Noise reduction

Temporal NR over 3 frames is good for subtle noise reduction in well exposed scenes. You generally want NR in any shot, and maybe add a bit of sharpening, because NR tends to soften the image.

Make use of Power windows in color correction, e.g. to gain down a bright region in shots where you had to illuminate the scene to bring more detail on a moving subject. Again, use temporal NR and sharpening to eliminate noise that’s even visible in low ISO.

In more extreme noise situations (indoors, high ISO), add spatial NR and possibly AI-based NR, which helps a lot.

### Transitions

Transitions to and from black are preferable over cross fades. They have a “render bypass” option to be encoded into the metadata track to render live on the headset. Cross dissolve on the other hand has to be baked into the pixels.

### Delivery settings

- **Bitrate**: you can increase the bitrate in the export of a Vision Pro Review. It wouln’t stream to AIV very well, but you can always AirDrop.
- **visionOS bypass**: “off” will bake in all effects; “transition settings” will follow transition settings on a case-by-case basis; “always” is recommended for best on-device render.
- **Audio**: choose the proper codec and bus, it should be ASAF/APAC.

### Export Presets

- **Vision Pro Review**: best for previewing on the AVP, but it currently uses the hardware encoder so the compression isn’t efficient. Resolve will add a better MV-HEVC software encoder/renderer soon.
- **Vision Pro Bundle**: best for archival/mezzanine file. The video is encoded in ProRes and everything is included at full resolution.
- **ASAF**: for exporting the audio master in only, in full resolution.

- **VR180:** for exporting the video in side-by-side half-equirectangular for other platforms. The quality will be degraded.

. . .

## The Elephant In the Room

The event delivered a treasure trove of information, and kudos to you if you’re still reading this.

But it only covered technical and creative aspects of immersive media. It didn’t touch at all on business considerations, which were top of mind for those present at the social mixer.

The unsaid assumption, throughout the sessions, was that all those videos would be made for consumers, some of them distributed by Apple TV, and the rest distributed via third party apps on the App Store.

The reality is that the consumer market on visionOS is too small to support any form of monetization with return on investment: it’s not even that

production costs are too expensive, it's that proceeds are going to be too small. Success is *thousands* of sales, even the most popular creators don't break even.

### Isn't it canceled anyways?

A persisting misconception outside our niche XR space is that the Apple Vision line is done. I was reassured that *this notion is absurd.* And while Apple won't simply issue a press release to deny rumors, they're leaving a trail of evidence:

1. A new M5 Vision Pro just hit the shelves, albeit quietly.
2. Apple's careers website advertises hundreds of Vision Pro openings.
3. Events like this one are signs of continued investment in the development of the ecosystem.
4. Expensive unannounced projects like the 100 immersive cameras at the Real Madrid stadium are still popping up.

### An elusive monetization

Apple famously doesn't invest in or give grants to startups, unlike Meta. Apple TV seems keen to license the top of the crop of immersive production, yet this won't be enough to bootstrap the critical mass of content that will make Vision Pro worth the price for consumers.

The thing is, Vision Pro was always for professionals and well-off early adopters; this is why the rumor of the shelving of a cheaper and lighter "Vision Air" is concerning. Even if the Pro line is here to stay, why make videos that very few people can experience?

Apple expects indie studios to lean on cross-platform distribution in order to maximize monetization. But AIV, as great as it is, remains a proprietary format that won't be supported on other devices anytime soon. Uncertainty and speculation surrounds Google and Meta's plans for XR in a world of

growing interest in smartglasses, and it doesn't look like the new Samsung Galaxy XR is going to make a difference, for now.

### Should we just wait?

There’s an inflection point some time in the future where the consumer market will be large enough to start supporting immersive video. We’ll even see immersive ads!

But it might take a couple years longer than initially thought. Creators I talked to during the event were following other business models:

- Those who can afford it or can raise money for it **produce videos as assets**, expecting to license them in several years once the market matures.
- Some produce for **ticketed location-based activations**, such as museums or retail, where watching immersive video on the Vision Pro is sold as a premium novelty experience.
- Others do B2B and enterprise projects, often in training and learning. Some brands apparently also angle for the “prestige” of a presence on such an exclusive platform.
- Relatedly, there is anecdotal opportunity in “premium wedding photography” type gigs for high net worth individuals.

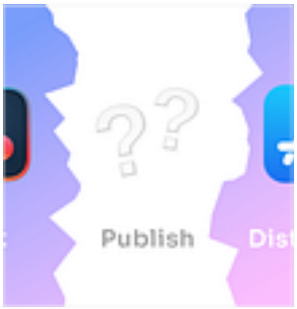
. . .

At Acute, we strive to help the community by making distribution easy, high quality, and affordable. A rising tide lifts all boats, and every new immersive video gets us a bit closer to the mainstream.

**You bought a $30,000 immersive camera, and now what do you do?**

With full support for Apple Immersive Video, Acute delivers on the filmmaker request: “I just want a big UPLOAD...

medium.com




Immersive Video | Apple | Apple Vision Pro | Virtual Reality

**Written by Anthony Maës**

264 followers · 10 following

Follow

Immersive Engineer & Founder, Acute Immersive https://acuteimmersive.com/

## Responses (1)

Write a response

What are your thoughts?

Iain Anderson
Nov 4, 2025

Really great! Thanks so much for taking and sharing these notes.

I'm really curious about a couple of things:

1. Will Apple promote Spatial as well as Immersive? I've been experimenting with 3D video and have created a plug-in for Final Cut Pro to... more

1 reply Reply

## More from Anthony Maës

Anthony Maës

## I wanted to love the Meta Ray-Ban Display glasses

Throughout the 10–15min demo I tried to squint and see what it’d be great at. I

Oct 9, 2025 128 1

Anthony Maës

## OpenImmersive 1.6: now playing side-by-side and over-under

A free (as in free beer and as in free speech) implementation of 3D VR180 and VR360

Jan 29 5 1

Anthony Maës

## Canon EOS R5 C & Dual Fisheye Lens Best Settings for Immersive

I fiddled around so you don't have to.

Jul 19, 2024 52 1

Anthony Maës

## Hacking HTTP Live Streaming with Swift

How to get AVFoundation to play a specific resolution/bitrate when Apple is trying to

Jun 14, 2025 3 1

See all from Anthony Maës

## Recommended from Medium

In Activated Thinker by Shane Collins

### Why the Smartest People in Tech Are Quietly Panicking Right Now

The water is rising fast, and your free version of ChatGPT is hiding the terrifying,

Feb 13 11.3K 496

In CodeX by MayhemCode

### Why Thousands Are Buying Mac Minis to Escape Issues with Big

Something strange happened in early 2026. Apple stores started running low on Mac

Feb 15 3.8K 65

In UX Planet by Nick Babich

## Claude Skills for Product Designers

How to create your first Claude Skill

Feb 10 537 7

In Mac O’Clock by Nov Tech

## Apple Just Fired the Designer Who Made iOS 26 Unreadable. Here’s

Now he’s going to Meta, Stephen Lemay is taking over, and Apple employees are

Dec 22, 2025 8.3K 313

In Write A Catalyst by Dr. Patricia Schmidt

## As a Neuroscientist, I Quit These 5 Morning Habits That Destroy Your

Most people do #1 within 10 minutes of waking (and it sabotages your entire day)

Jan 15 33K 624

Adam Dev

## 🌊 iOS 26 Introduces the Liquid Tab Bar

A new era of fluid navigation in SwiftUI

Dec 12, 2025 3

See more recommendations