

ONE YEAR WITH APPLE IMMERSIVE VIDEO

What the medium can really do – and where it is going

Insights from a year of production, many screenings, and direct exchange with Apple



Carl Amadeus Hiller

Apple Immersive Video

Practical Experience and Insights

Contents

1	Presence, temporal fidelity, and authenticity	3
2	Why an impressive location is not enough	12
3	We Make Belonging: people, architecture, and social space	14
4	Architecture as a field in its own right	24
5	Documentary work in real living environments	27
6	Music, performance, and spatial sound	34
7	Film production, workflow, and fictional storytelling	37
8	The areas where the medium is strongest	46
9	The limits of the medium	48
10	How Apple Immersive Video is developing	49
11	Museums, showrooms, and physical screening spaces	53
12	The economic reality	55
13	Positioning and conclusion	57

Insights from practical work, many screenings, and direct exchange with Apple

After my experiences with *We Make Belonging*, *Vivaldi Immersive*, and many other documentary, artistic, and music productions - from working with fishermen on a small Aegean island, to portraying a violin maker, to close encounters with exceptional musicians - as well as many screenings and direct exchange with Apple, I can now describe Apple Immersive Video much more clearly:

Apple Immersive Video is a medium for a physical sense of presence. Its greatest strength lies in the connection between people, action, and space.

It is not simply about recording a location in 3D. It is also not about technically upgrading a normal film. A project works especially well when the audience experiences something through its presence that would feel fundamentally different or weaker on a flat screen.

For me, the key question is therefore:

Why does the audience need to be present inside this particular story?

Apple describes presence, authenticity, closeness, and connection as the creative foundations of spatial and immersive storytelling. The immersive version of a project should therefore not be only an additional technical format. It needs its own storytelling value and, ideally, should be the strongest possible form of that story.

My productions have led me to another conclusion:

A meaningful location is another major strength.

By this, I do not simply mean an attractive location. I mean a space that has been designed, is changing, brings people together, or makes an action understandable in the first place.

Carl Amadeus Hiller, July 2026



The effect of Apple Immersive Video is difficult to describe in words. The key is the experience of perceiving a person and a space in a physically believable way.

01

Apple Immersive Video does not simply show 3D - it creates the feeling of having been there

The special quality of Apple Immersive Video does not come from its enormous resolution alone. What matters is the combination of:

- stereoscopic depth,
- believable spatial scale,
- a very wide field of view,
- high temporal resolution,
- and sound that remains stable in space.

Apple Immersive Video was designed to capture the real world with a high level of spatial and temporal accuracy. Calibration, projection, and motion data are preserved through the format until playback on Apple Vision Pro. The audience should not only see a person, object, or room as large and detailed, but experience it at a believable size and distance.

This is a fundamental difference from conventional film. In a two-dimensional film, a wide shot can look magnificent, but it remains an image. In Apple Immersive Video, the viewer is placed in a physical relationship with the space:

- Am I inside or outside?
- Is a person standing two meters away or ten meters away?
- Does the room feel intimate or monumental?
- Do I feel welcome, exposed, protected, or overwhelmed?
- Can I truly understand the size of a building, a work of art, or a group of people?

For this reason, immersive scenes can remain in the memory more like a personal experience than like a film that has been watched.

I do not only remember what I saw. I remember a spatial situation, almost as if I had been there myself.

Why the brain accepts the image as so believable

The unusual authenticity of Apple Immersive Video does not come from resolution, stereoscopy, or the wide field of view alone. The temporal accuracy of the recording is just as important.

The human brain does not recognize other people only through static facial features. It is extremely sensitive to biological movement and to the timing of movement in a face and body.

Before a person speaks, smiles, or visibly reacts, very small changes have already begun:

- The lips and cheeks tighten.
- The eyes shift.
- The head moves slightly.
- The person breathes in.
- Individual facial muscles begin to activate.
- Posture and muscle tension change.

Some of these movements take place within very short periods of time. Very brief involuntary facial expressions are often called micro-expressions. But ordinary communication also consists of many small phases of movement: the beginning of a reaction, its acceleration, its peak, and the way it fades.

The brain processes these patterns largely without conscious effort. It is very good at recognizing whether a face, body, or gesture is moving naturally. Facial expressions can feel less convincing and more artificial when their natural timing is changed.

This is where recording at 90 frames per second becomes important.

At 60 frames per second, movement is sampled every 16.67 milliseconds. At 90 frames per second, a new image is created every 11.11 milliseconds. This provides 50 percent more temporal information for the same movement.

A short movement lasting 100 milliseconds is described by about six individual frames at 60 frames per second. At 90 frames per second, it is described by nine. The beginning, acceleration, peak, and fading of

a small movement can therefore be recorded and reproduced in greater detail.

This does not mean that 60 frames per second cannot show these reactions at all. There is no single hard threshold at which perception suddenly changes. But at 90 frames per second, natural movement is described more accurately. Motion blur, strobing, and visible jumps between individual frames are also reduced.

This affects more than faces. It also affects:

- the movement of hands,
- breathing,
- clothing and hair,
- a bow moving across a violin string,
- the vibration of an instrument,
- a working movement in a workshop,
- or the physical energy of a musician.

The result is greater temporal fidelity in biological movement.

More signals now agree with one another for the brain:

- Spatial depth feels believable.
- People and objects have a natural scale.
- Movement flows continuously over time.
- Face, body, and voice fit together.
- Sound remains anchored in the room.
- The image responds immediately when the viewer moves their head.

The brain therefore receives fewer signals that it is looking at a technical reproduction. It processes the situation more like a real encounter.

This helps explain part of the difference between Apple Immersive Video and many previous VR180 experiences. It is not only the number of pixels. It is also not only the field of view. What matters is the combination of spatial and temporal accuracy.

Apple Immersive Video does not only record what a person looks like. It preserves more of how a real person behaves over time.

These fine details are often not noticed consciously. But they play a major role in making a person feel less like a three-dimensional video image and more like someone who is actually standing in the room.

Authenticity becomes much harder to fake

The high spatial and temporal accuracy has another consequence: Apple Immersive Video quickly reveals whether a person is believable in what they are doing.

This applies both to documentary subjects and to actors in fictional productions.

In conventional two-dimensional film, a scene can be controlled strongly through selection and editing:

- A reaction can come from a different take.
- Uncertainty can be hidden with a reverse shot.
- An action can be built from several short shots.
- Framing, focus, and editing guide the audience's attention very precisely.
- Small physical contradictions can remain outside the selected frame.

In Apple Immersive Video, this is only possible to a limited degree. A person often remains visible for longer and with more of their body in view. The audience can choose whether to watch the face, the hands, the posture, or a small movement in the background.

This makes unconscious signals easier to notice:

- an uncertain look,
- a short hesitation,
- body tension that does not fit the situation,
- an action performed mechanically,
- nervous movements of the hands or feet,
- a contradiction between spoken words and physical behavior,
- or an emotion that is shown only in the face but not supported by the rest of the body.

This does not mean that fictional acting must feel artificial. On the contrary, a strong performance can be extraordinarily intense in Apple Immersive Video. But it must remain believable through the face, voice, breathing, posture, movement, and the person's behavior in space.

The real issue is therefore not simply the difference between "real" and "acted." It is about truthful behavior.

An actor can experience a completely fictional situation in a believable way. A documentary subject, on the other hand, can become unconvincing as soon as they begin to perform a version of themselves for the camera.

Because of the high frame rate and the visible micro- movements, the audience receives a great deal of information about how a person is actually behaving within a moment. These signals come not only from a selected close-up, but from the whole visible body and the surrounding space. Inconsistencies are therefore more difficult to hide.

Apple Immersive Video requires especially careful casting and selection of the people in front of the camera.

For documentary projects, I need people who feel safe enough with me and with the camera to stop performing a role. Trust, time, and a calm filming situation become even more important.

For fictional productions, I need actors who can do more than create a performance for individual shots. They must carry a situation physically, emotionally, and spatially over a longer period of time.

The more believably the medium reproduces a person, the more believable that person's behavior must be.

Apple Immersive Video therefore does not only strengthen authenticity. It also exposes a lack of authenticity more quickly.

This is a major opportunity for both documentary and fictional storytelling. But it also means that casting, trust, preparation, and directing actors have an even stronger influence on the result than in many conventional production formats.

Encountering your own spatial image

When Morag Myerscough saw herself in *We Make Belonging* on Apple Vision Pro, she was almost shocked at first. She described the feeling as meeting her own digital twin.

This reaction is understandable. We normally see ourselves either in a mirror or in a two-dimensional photograph or video.

A mirror shows a reversed image from a familiar perspective. A conventional video gives us an outside view, but it reduces our body and the surrounding space to a flat image. Spatial depth, real distances, physical scale, and much of the relationship between the person and the environment are lost.

Something different happens in Apple Immersive Video.

A person encounters themselves:

- at a believable physical size,
- with spatial depth,
- inside the real environment,
- with their own posture and body tension,
- with the movement of their hands, face, and eyes,
- with their distance from other people,
- and with the effect they have within the room.

You do not only see what you look like. You experience from the outside how you behave in a space.

The high temporal and spatial accuracy can make this encounter unusually direct. Small movements that we normally do not notice in ourselves become visible:

- a short hesitation before a sentence,
- our own breathing,
- repeated gestures,
- unconscious tension,
- the position of the shoulders,
- eye movements,
- the distance from an instrument or conversation partner,

- or the way we enter a room and take up space within it.

This creates another possible field of use for Apple Immersive Video:

Self-observation, learning review, coaching, and reflection.

For musicians, this could make the following aspects visible and audible:

- posture and breathing,
- economy of movement,
- distance from the instrument,
- interaction with other musicians,
- spatial balance within an ensemble,
- stage presence,
- and the effect of the performance on a viewer in the room.

A musician would not only review the performance as picture and sound. They could encounter themselves almost like another person in the real performance space.

This kind of self-observation could also be valuable for presentations, lectures, and teaching situations:

- How do I stand in the room?
- Do I use the space consciously, or do I remain in one place?
- Where do I look?
- How do my gestures appear?
- Do I make myself look physically large or small?
- How close do I come to other people?
- What effect do pace, pauses, and movement have?
- Does my physical presence support what I am saying?

The same applies to acting, dance, sport, conducting, vocational training, and other activities in which not only the result but the entire physical and spatial process is important.

Apple Immersive Video should not be seen as an objective assessment machine. A person's view of their own recording remains subjective. The medium does not replace teachers, coaches, directors, or outside

observers. But it can provide a much richer basis for shared analysis and reflection.

It has not yet been clearly established whether immersive video is generally more effective as a learning tool than conventional video. At present, this is a well-founded possibility, not a proven superiority.

Even so, the potential is significant:

Apple Immersive Video can be more than a window into other people and places. It can become an unusually precise mirror of our own impact.

Encountering your own spatial image can initially feel unsettling or even frightening, as it did for Morag. But this is also where its possible value lies: you do not see yourself as you imagine yourself internally. You experience how your body, movement, and presence appear from the outside in a real space.



Morag Myerscough encountering her own spatial image. Immersive playback can feel like meeting a digital twin.

02

Why an impressive location is not enough

An important insight concerns the idea of teleportation.

At first, it seems obvious that the greatest strength of the medium is its ability to transport people instantly to extraordinary places. This can be very impressive at the beginning. But even a spectacular landscape or an exceptional building can eventually feel like a very high-quality spatial postcard.

A location becomes a story only when I meet a person there, witness something happening, or build a relationship with someone and their activity.

This applies to almost every genre:

- A beautiful landscape is not yet a story.
- An exceptional building is not yet a story.
- A stage full of musicians is not yet a story.
- A famous person in front of the camera is not yet a story.

The medium becomes powerful through the combination of:

Person + activity + place + change.

This connection runs through all of my previous work.

With the fishermen on Fourni, it is not only the Aegean Sea, but their life and work within that landscape.

With the violin maker, it is not only the workshop, but the focused relationship between person, material, tools, and sound.

With Takuya, it is not only a musical performance, but the physical presence of a musician, his energy, and his immediate relationship with the instrument.

With Morag Myerscough, it is not only a colorful installation, but the creation of a social space by an artist and many participating people.



Morag Myerscough in her studio. In *We Make Belonging*, person, process, and space were considered together from the very beginning.

03

Why “We Make Belonging” is such a fitting Apple Immersive Video project for us

Two different perspectives on people and space

I developed and made *We Make Belonging* together with my partner, Isabelle Hiller. Isabelle is an interior architect. This was especially important for the film.

We did not look at Morag Myerscough's work only from a filmmaking perspective. At the same time, we explored the effects of architecture, color, material, proportion, movement, and social space.

Isabelle reads spaces differently from a conventional film crew. For her, a room is not simply a background in which an action takes place. She looks at:

- how people are guided through a space,
- how sightlines develop,
- how color and light change perception,
- how materials respond to one another,
- which areas create closeness or distance,
- how architecture influences behavior,
- and what turns an empty area into a place where people want to stay.

My own perspective came more strongly from the camera, documentary observation, and the question of how we could make Morag and the creative process feel alive on film.

The combination of these two perspectives was essential to the project:

We did not simply film inside an exceptional space. We tried to understand how the space works, how it changes, and what it does to the people inside it.

The subject and the medium fit together directly

Morag Myerscough transforms public spaces through color, patterns, language, structures, and community participation. Her work deals deeply with the question of how spaces can become places where people feel belonging, joy, and community.

Apple Immersive Video does not simply record this work. The medium continues it on another level.

The audience does not view Morag's work from the outside. It is placed:

- inside her world of color and form,
- between different spatial layers,
- in relation to the people working there,
- close to Morag during her creative process,
- and inside a place that is visibly changing.

This is a central point for us:

Our film is not only about immersive design. It allows the audience to experience that design immersively.

With many subjects, we first have to explain why they should be told in Apple Immersive Video. With Morag's work, the reason is almost self-evident.

Architecture is not only shown - it is physically experienced

In traditional architectural photography and two-dimensional film, a space can be explained visually:

- through wide shots,
- camera movements,
- floor plans,
- details,
- and cuts between different perspectives.

But the audience remains outside the room. It has to build a mental understanding of the spatial relationships from several flat images.

Something different happens in Apple Immersive Video. The viewer is given a fixed physical position. From that position, they experience at the same time:

- foreground, middle ground, and background,
- height and depth,
- views through spaces and overlapping layers,
- scale,
- material qualities,
- light,
- color in the peripheral field of view,
- people as a measure of scale,
- and the acoustic character of the location.

Morag's work benefits enormously from this. Her colors and geometric forms do not exist only inside a rectangular image. They surround the audience. Words and patterns can be discovered independently. The different elements do not compete for space inside a limited frame, but form a complete spatial composition.

Isabelle's architectural perspective was especially valuable here. It helped us look not only for a spectacular image, but for a camera position from which the whole space could be understood.

We had to consider:

- What is in front of the viewer?
- Which architectural layers lie behind it?
- How do color and structure work in the peripheral field of view?
- Where are the people in relation to the installation?
- Which spatial change becomes understandable from this position?
- Does the viewer feel like an observer or part of the situation?

This is one reason why the film works so well: the audience does not only understand intellectually that Morag changes spaces. It feels the change.

The space is not a background - it is a main character

In *We Make Belonging*, there are really three main characters for us:

1. Morag Myerscough,
2. the people working with her,
3. and the changing space itself.

FLUX is not a decorative background. It connects the museum, the city, and the community. Morag's work is designed not only to change public spaces visually, but to give them a new identity through color, language, and participation.



The camera inside the developing FLUX space. The film is not only about the finished work, but also about the spatial transformation.

Our film therefore follows more than the creation of an artwork. It shows the creation of belonging:

- People develop words and patterns.
- They paint and build together.
- Morag responds to the existing architecture.
- The way the place is perceived changes.
- An empty spatial gap becomes a social place.

In a conventional architecture film, we could document the final result. In Apple Immersive Video, we can also make it possible to experience how a spatial idea takes over a place.

The audience is present inside this change. It does not only watch a space being created. For a moment, it becomes part of that process.

The human anchor prevents the postcard effect

The film does not work only because FLUX is large, colorful, and visually striking.

It works because Morag guides us through these spaces as a person:

- through her attitude,
- her concentration,
- her energy,
- her decisions,
- her hands at work,
- her relationship with Luke,
- her conversations with the people involved,
- and the connection between her life in London and the project in Munich.

The audience experiences more than an installation. It understands where the work comes from and why it matters to Morag.

Morag gives the architecture emotional meaning. Without her, FLUX might be an exceptional spatial recording. With her, it becomes a story about work, identity, community, and belonging.

For us, this is the difference between a technically impressive recording and a truly successful immersive film.

A real process is stronger than a staged explanation

Another important lesson from the production is that Apple Immersive Video becomes especially strong when people are actually doing something.

Morag paints, designs, discusses, decides, and works. Other participants build and paint real surfaces. These are not reconstructed actions performed only for the camera.

In a high-resolution spatial image, artificial actions and unconvincing behavior are easier to notice. The audience can study details, tools, movement, and body language. Real skills and genuine processes therefore usually work better than constructed activity.

At times, our way of working felt more like the careful approach of shooting on analog film than an ordinary digital documentary shoot:

- We had to observe very closely beforehand.
- We had to read the space.
- We had to wait for the right moment.
- We had to prepare the process carefully.
- We had to place the camera very consciously.
- Then the situation had to work in a longer, precise take.

The format forces us to concentrate. We cannot record every possible direction and repair everything later through fast editing. At the same time, this creates a calmness and honesty that fits Morag's creative process very well.

The audience is allowed to discover things for itself

In conventional film, we largely control where the audience looks.

In Apple Immersive Video, we give up part of that control. Viewers can make their own discoveries within the visible space:

- read a sentence on a wall,
- watch someone painting,
- follow Morag,
- examine the architecture,
- notice details in the distance,
- or discover a moment that we did not emphasize through a cut.

This freedom is not a technical side effect. It is part of immersive storytelling.

Instead of forcing the audience's attention through pans or rapid editing, we can guide it through composition, movement within the space, light, and Spatial Audio.

Morag's work is particularly suited to this because her spaces contain many layers that can be read at the same time. The audience can find its orientation and still continue to discover new details.

The same film can therefore feel different on a second viewing. Not because the film has changed, but because viewers place their attention differently.

Belonging is not only explained - it is communicated through space

The title *We Make Belonging* describes more than Morag's philosophy. It also describes what the medium itself can do.

A two-dimensional film can explain that a place was designed to be open, inclusive, and communal. In Apple Immersive Video, the audience can experience for itself:

- whether a space feels generous or narrow,
- whether people move through it freely,
- whether you feel part of the situation,
- and whether the architecture creates distance or closeness.

This is a special form of social impact. It does not come only from an argument, but from a spatial experience.

With subjects such as:

- belonging,
- public space,
- inclusion,
- community,
- cultural participation,
- urban development,
- and identity,

Apple Immersive Video can communicate something that is difficult to describe in words alone.

This may be one of the main reasons why *We Make Belonging* has such an immediate effect in screenings. People do not first talk about pixels, resolution, or stereoscopy. They respond to color, closeness, people, space, and atmosphere.

The technology disappears behind the experience. For us, that is a sign that an immersive film is working.



Filming during the construction of FLUX. Camera, architecture, and real work come together.



Isabelle Hiller and Carl Amadeus Hiller with the URSA Cine Immersive on location.

04

Why architecture is an important field for Apple Immersive Video

Architecture is not automatically a good immersive subject. A simple 180-degree view of an empty building can become just as tiring as a landscape without a story.

Architecture becomes especially powerful in Apple Immersive Video when at least one of the following conditions is present.

The space is part of the action

This is true of FLUX. The story is about how a space is created and how it changes people.

Other possible examples include:

- the construction of an installation,
- the renovation of a historic building,
- a stage set while it is being created,
- sacred architecture and the rituals that take place inside it,
- an artist's studio or workshop,
- a special concert hall,
- a school,
- a hospital,
- or a public place with a social purpose.

The location is not normally accessible

Apple Immersive Video can take people into spaces they would not normally be able to enter:

- backstage areas,
- private artists' studios,
- historic buildings,
- production facilities,
- architecturally important private homes,
- construction sites,
- endangered locations,
- or temporary installations that exist only for a short time.

The space will disappear or change

This creates another important function: Apple Immersive Video can become a spatial memory.

An installation, a stage design, or a temporary structure can be preserved as more than a photographic record. The place can remain available in its real scale, atmosphere, and relationship to the human body.

With *We Make Belonging*, we are therefore not only documenting FLUX. We are also preserving a particular stage of its creation:

- the people involved,
- the unfinished states,
- the working processes,
- and the atmosphere of that period.

A person gives us access to the space

The safest way to protect architecture from becoming a postcard is to include a believable person:

- an architect,
- an artist,
- a resident,

- a craftsperson,
- a historian,
- a musician,
- or someone whose life is closely connected to the place.

The person does not have to explain the space all the time. Often, it is enough to see them working and responding to it. The audience then understands the place through the relationship between the person and the architecture.

05

Documentary work: people inside their real living environments

My documentary productions have shown me that Apple Immersive Video is especially strong when a person's environment is not only a background, but part of their identity.

The fishermen on Fourni

Faces of Fourni: Bound to the Sea is not simply about beautiful images of a Greek island.

The landscape, the sea, the boat, and the physical work are all part of the two fishermen's lives. The space explains their work, and their work explains the space.

The strength comes from the combination of:

- the openness of the sea,
- the limited space on the boat,
- the distances between the people,
- the physical work,
- the weather,
- the sounds,
- and the fishermen's relationship with their surroundings.

The camera does not only take the audience to an island. It takes them to a real workplace and into a lived reality.

Without the people, the Aegean Sea would be an impressive travel destination. Through the fishermen, it becomes a living environment.



Filming on Fourni. The landscape is not presented as a postcard, but as a place of life and work.

The violin maker

The portrait of a violin maker uses an almost opposite spatial situation.

Instead of openness, there is concentration and closeness:

- a workbench,
- tools,
- wood,
- hands,
- fine movements,
- material surfaces,
- and a sound that is slowly coming into being.

The audience does not only see the finished instrument. It experiences the spatial relationship between the craftsman, the material, and the tools.

Apple Immersive Video is especially strong with craft processes. The work does not need to be dramatized through many cuts. The viewer can observe how someone works, how confident their movements are, and how much experience is present in small actions.

The room becomes an extension of the person. The workshop already tells us something about the person before they explain anything.



The violin maker at work. Closeness, material, and fine movements become especially clear in the immersive format.

Takuya and the physical experience of music

Working with the Japanese drummer Takuya reveals another strength: the immediate physical presence of a musician.

In Apple Immersive Video, music is not only something we hear. It becomes visible as movement in space and can be understood physically.

This includes:

- posture,
- movement of the hands and arms,
- distance from the instrument,
- strength and precision,
- concentration,
- and the way sound spreads through the room.

The musician is not shown only as a performer. The audience enters a direct relationship with his body, his instrument, and his energy.

Drums and percussion have a particularly strong spatial and physical dimension. Movements are large, sound sources can be clearly located, and rhythm becomes visible.

But the distance must be chosen carefully. The closeness should feel intense without becoming uncomfortable. Finding this balance is one of the new creative tasks of the medium.



Takuya performing. Body, instrument, and space become one connected experience.



Vivaldi Immersive: The baroque musician wearing Apple Vision Pro. Historical performance practice meets a new spatial medium.

06

Music, performance, and spatial sound

For me, music is one of the most convincing fields for Apple Immersive Video.

A conventional concert recording normally shows a stage from changing camera angles. In an immersive film, a different question becomes important:

Where can I place the audience that they could never be during a normal concert?

This can mean:

- very close to a musician,
- inside an ensemble,
- in a rehearsal room,
- in a special architectural space,
- during a concentrated moment before the performance,
- or in a position normally reserved for the musicians themselves.

Vivaldi Immersive shows me how powerful the combination of musical performance, closeness, and spatial sound can be.

But such a film does not work through technical perfection alone. Music productions also need:

- different degrees of closeness,
- spatial development,
- a dramatic progression,
- a relationship with the musicians,
- and careful control of attention.

Spatial Audio is not simply an added improvement to the soundtrack. It is a basic part of the experience.

The sound should not feel trapped inside the headphones. It should remain believably located inside the recorded space. When the viewer moves their head, the acoustic world must remain stable.

Music and architecture meet here directly:

The space itself becomes an instrument, and the music makes its dimensions audible.



The URSA Cine Immersive inside the church. Architecture and sound become part of the story together.



The URSA Cine Immersive in the violin maker's workshop. Despite the extraordinary effect, the core process remains filmmaking.

07

A major advantage: Apple Immersive Video remains film production at its core

An important point is often overlooked in discussions about immersive media:

A linear Apple Immersive film is not a fully computer-generated or interactive XR application.

I do not have to build a virtual world from the ground up. For a conventional documentary, artistic, or music project in Apple Immersive Video, I normally do not need:

- a game engine,
- complex real-time programming,
- fully modeled 3D environments,
- volumetrically captured actors,
- a detailed photogrammetry model of the entire location,
- or programs such as Blender, Unity, or Unreal Engine.

These tools can be useful when I want to add interactive elements, elaborate computer-generated worlds, spatial graphics, or additional 3D objects. But they are not a basic requirement for a linear Apple Immersive film.

Apple designed the AIV system around a traditional film workflow:

***All-in-one camera → single file → traditional edit
→ encode → delivery.***

The camera records both eyes, optical calibration, and other necessary metadata in one connected file. The technical complexity remains below the surface and is managed by tools that support Apple Immersive Video.

With the URSA Cine Immersive, both sensors are recorded into a single Blackmagic RAW file. There is no complicated stereo camera rig whose two cameras must later be synchronized manually, geometrically corrected, and combined.

The footage can be loaded directly into DaVinci Resolve. Editing, color grading, selected VFX work, Spatial Audio, and output all take place within one software environment.

This is a fundamental difference from many earlier VR, XR, and 3D production methods.

The effort is comparable to a high-end 2D production - but distributed differently

Shooting an Apple Immersive Video project does not automatically mean creating a huge visual-effects production.

At its core, I need the same basic skills and production stages as I would for a high-quality two-dimensional film:

- a strong idea,
- clear storytelling,
- suitable subjects or performers,
- research,
- production planning,
- camera,
- lighting,
- sound,
- editing,
- color grading,

- sound mixing,
- and careful finishing.

Of course, there are major differences.

The data volumes are much larger. Recording takes place at extremely high resolution and high frame rates. Rendering time, storage, and technical quality control are more demanding. Sound has to be planned and produced spatially.

Visual design also requires new decisions:

- The camera sees almost the entire space in front of it.
- The crew is difficult to hide.
- Lighting, sound, and production design must work across a very wide field of view.
- The distance from people is critical.
- Fast camera movement and restless editing can feel uncomfortable.
- Shots need to be checked regularly on Apple Vision Pro.
- The audience can decide for itself where to look.

The main differences are therefore less about entering a completely new software world and more about spatial design, guiding the viewer, and handling the technical size of the material.

A faster path from recording to a finished experience

This simplification is one of the greatest advantages of Apple Immersive Video for me.

In a fully interactive XR production, a digital world often has to be created first:

- Spaces are modeled.
- Materials are built.
- Lighting systems are set up.

- Characters are animated.
- Interactions are programmed.
- User interfaces are developed.
- The entire experience is optimized for a real-time engine.

In a documentary Apple Immersive project, I can instead record a real person in a real place doing a real activity and then work with the result in a largely conventional post-production environment.

This allows me to move much more directly from the idea to the finished experience.

Creative energy does not have to go first into building a technical simulation. It can focus on what I already understand as a filmmaker:

- observing people,
- preparing situations,
- finding the right camera position,
- shaping light and space,
- developing a story,
- and creating an emotional rhythm.

This makes Apple Immersive Video especially interesting for:

- documentary film,
- art,
- architecture,
- music,
- performance,
- craft,
- portraits,
- culture,
- nature,
- and real working processes.

I can create the presence of an XR experience without necessarily constructing a complex XR world.

Simpler technology does not make the creative work simple

This point should not be misunderstood.

A familiar technical workflow does not mean that Apple Immersive Video is simply "2D with two lenses."

The special technical processes may be handled largely by the camera, metadata, and software. But creative responsibility does not become smaller. It shifts.

The key questions are:

- Where should the camera stand?
- How close can I place it to a person?
- What is happening at the same time across the full field of view?
- Where is the audience looking before and after a cut?
- How long must a shot last for the space to become understandable?
- How do I guide attention without forcing it?
- Which objects or crew members will become visible?
- How does a person's effect change through their real distance from the camera?
- Where is the acoustic focus?

Apple Immersive Video has become more technically accessible, but creatively it remains a new medium.

This is precisely what makes it exciting: the tools are related to conventional filmmaking, but the storytelling decisions have to be learned again.



Fictional filming in a public space. Apple Immersive Video requires spatial choreography that goes beyond individual shots.

Fictional storytelling: thinking in scene arcs and spatial action

For fictional productions, this form of perception creates another fundamental change.

Apple Immersive Video can be edited, but cuts are more demanding than in conventional film. After a cut, the audience has to understand not only what has changed. It must also reorient its gaze and adjust to a possibly different viewing distance.

The audience also needs time to explore a space. A cut cannot always happen immediately after the central line of dialogue. The viewer may be looking at another person, noticing a detail in the room, or following a movement that the director cannot fully control.

Longer shots and continuous action therefore become more important.

The director has to think more strongly in spatial and temporal arcs:

- How does a scene develop without constant cutting between close-ups and reverse shots?
- How do the characters move in relation to one another and to the audience?
- When does a person enter the visible space?
- Where does new attention appear?
- How does focus move through action, light, movement, or sound?
- What happens while another character is speaking?
- What physical development does an actor go through during the full shot?
- How does the scene end in a way that prepares a possible cut?

In this sense, the work becomes partly closer to theater.

As on a stage, actions, movement, entrances, exits, and relationships between characters have to work over a longer period of time. Actors can rely less on having their performance assembled later from many individual moments. They have to carry a scene as one continuous development.

The director also cannot think only in isolated shots or individual reactions. A complete emotional and spatial arc has to be developed. Even so, Apple Immersive Video is not simply filmed theater.

The camera position gives the audience a very specific place inside the scene. Closeness has a different intensity than it does on a stage. Spatial Audio can guide the viewer's attention. Movement can approach the viewer or surround them. Carefully prepared visual connections still make powerful cuts possible.

The language therefore lies somewhere between film and theater:

- From film, it takes camera, editing, light, sound, and control over time.
- From theater, it takes the importance of blocking, ensemble, spatial continuity, and a performance sustained over a longer period.

For fictional projects, much more therefore has to be decided and tested before filming:

- casting,
- rehearsals,
- blocking,
- production design,
- sightlines,
- movement patterns,
- spatial storytelling,
- possible cut points,
- and the audience's position within the action.

Fictional Apple Immersive Video requires less thinking in individual images and more thinking in scenes that can be experienced.

The director is not only staging what the camera shows at a particular moment. They are creating a space in which the audience remains present with the characters over a longer period of time.



Performance and spatial choreography. Movements, pathways, and sightlines must work across longer sequences.

08

Where Apple Immersive Video is especially strong

From my experience so far, six areas are especially promising.

People in meaningful spaces

Artists, architects, craftspeople, witnesses to history, or exceptional personalities in a place closely connected to their work and identity.

Processes that can be observed

Building, designing, making music, training, repairing, rehearsing, researching, or working. Activities in which the body, tools, materials, and spatial processes are important.

Performances and events

Concerts, sport, theater, dance, and other events where there is a clear value in being present at an unusual position.

Inaccessible or temporary places

Locations that are far away, dangerous, exclusive, available only briefly, or likely to change later.

Spaces with social or cultural importance

Museums, public architecture, memorial sites, cultural meeting places, and projects that make social relationships understandable through space.

Education, training, and professional self-reflection

Musicians, actors, speakers, teachers, dancers, athletes, or trainees can observe themselves inside a real spatial situation.

This makes it possible to analyze not only individual technical mistakes, but also posture, presence, movement, timing, eye contact, and the relationship with the surroundings.

Apple Immersive Video could offer a new form of video feedback: not only watching a flat image, but encountering your own spatial image.

09

Where the medium is less suitable

Not every strong subject becomes better through Apple Immersive Video.

The following are often less effective:

- ordinary interviews without a meaningful spatial situation,
- empty architecture without people or change,
- random landscape footage,
- content that depends mainly on fast editing,
- concert recordings from one permanently unchanged audience position,
- product presentations that are only intended to appear larger and more three-dimensional,
- and projects whose main goal is to reach the largest possible audience in the shortest possible time.

Apple Immersive Video is also not a freely walkable digital twin. The audience can look around, but it remains at the camera position and inside a film guided through time.

Its strength therefore does not lie in complete spatial freedom, but in a:

***carefully designed encounter with a real person,
a real process, and a real place.***

10

How Apple Immersive Video is developing

From a technical demonstration to a production system

In the first phase, Apple's own productions mainly demonstrated what Apple Vision Pro and the medium could do.

This is increasingly developing into a complete production ecosystem:

- purpose-built cameras,
- integrated calibration and metadata workflows,
- immersive post-production,
- Apple Spatial Audio,
- specialized encoding methods,
- streaming,
- app-based distribution,
- local quality control,
- and external hosting and distribution services.

This is an important stage of development. Apple Immersive Video is no longer only a format for individual films produced by Apple itself.

It is becoming a system in which producers, rights holders, institutions, and platforms can create and distribute their own content.

From individual films to more complex forms of production

At the same time, the development is moving toward:

- multiple camera perspectives,
- live events,
- real-time monitoring,
- dedicated immersive vision mixing and direction,

- live Spatial Audio production,
- and combinations of filmed material with additional spatial elements.

This does not mean that smaller documentary films will become less important. On the contrary, they often show especially clearly what closeness, authenticity, and presence can achieve.

At the same time, production forms are emerging that are closer to professional broadcasting, live concerts, and sports coverage.

The medium is beginning to divide into different fields and formats.

From entertainment to a wider range of uses

Apple Immersive Video is developing beyond conventional entertainment into other areas:

- training and education,
- architecture and design,
- corporate communication,
- art and culture,
- museums,
- tourism,
- showrooms,
- documentary work,
- coaching and presentation training,
- music and acting education,
- movement and performance analysis,
- self-observation and learning review,
- and individually developed applications.

This does not mean that every corporate film or every museum automatically needs an immersive production.

Real added value appears only when at least one of the following is important:

- understanding space,
- realistic scale,
- personal closeness,

- precise observation,
- access to a special place,
- the physical experience of a process,
- or spatial reflection on your own presence and impact.



Screenings and guided presentation are part of the current reality of the medium. Many people need to experience Apple Immersive Video before they can truly understand its value.

11

The role of museums, showrooms, and physical screening spaces

I would not say that showrooms are the main market for Apple Immersive Video.

A more accurate statement is:

Showrooms, museums, festivals, and other supervised screening spaces are currently one of the most practical ways to use the medium.

This has less to do with the content itself than with the current market situation:

- The number of private Apple Vision Pro owners is still limited.
- The hardware is expensive.
- Many people need to experience the medium before they understand its value.
- Institutions can provide several devices in one central location.
- The screening context can be controlled.
- A specific audience is already present at the location.

This is especially interesting for *We Make Belonging*. The film can be shown in a museum, at a festival, or close to FLUX itself.

This creates a double experience:

- The audience experiences Morag, the creative process, and the space immersively.
- It can then encounter the real place itself.

This is more than a film screening. It combines:

- documentary,
- exhibition,
- architecture,
- and a physical visit.

The same applies to:

- concert halls,
- architecture centers,
- tourism locations,
- brand spaces,
- art fairs,
- company headquarters,
- and educational institutions.

The showroom is therefore not a replacement for wider future distribution. At the moment, it is a useful way to place the medium where there is a clear value and a specific audience.

12

The economic reality

The technical development of the medium is ahead of the mass market.

Productions in Apple Immersive Video remain demanding:

- The data volumes are enormous.
- Rendering times can be long.
- Grading and retouching are demanding.
- Spatial Audio requires its own production process.
- Encoding and hosting must be planned.
- The number of reachable end devices is still limited.

At the same time, this effort should not be confused with a fully computer-generated XR production.

The shoot and basic post-production of a linear Apple Immersive Video film can remain organizationally very close to a high-quality 2D film production. The additional work mainly comes from:

- data volume,
- spatial design,
- technical quality control,
- Spatial Audio,
- and the more demanding output process.

A virtual world does not first have to be developed and programmed over many months.

This makes it possible to reach a finished, high-quality experience relatively quickly - especially when real people, real places, and documentary processes are at the center of the project.

Even so, the business model of producing a film entirely at your own risk and then selling a very large number of individual tickets remains difficult to calculate at present.

More realistic models currently include:

- commissioned productions,
- projects financed by institutions or brands,
- cultural funding,
- licensing,
- screenings at specific locations,
- festival and museum presentations,
- dedicated apps from rights holders,
- and productions with parallel 2D or VR180 versions.

A pay-to-own model could become interesting for exclusive music, art, or collector's productions. But it is still unclear whether this can become a reliable business model.

The technological quality of the medium is already exceptional. Distribution and business models, however, are still at an early stage.

13

What my work so far also demonstrates

My position is not based only on the fact that I began working with professional immersive camera technology at an early stage.

I combine several skills that are essential for this medium:

- documentary observation,
- experience with artists and musicians,
- an understanding of architecture and space,
- conventional camera and lighting design,
- patience with real processes,
- experience in music production,
- complex immersive post-production,
- and a sense of when the technology has to disappear behind the story.

The different projects each demonstrate different strengths:

- *We Make Belonging* represents architecture, art, color, participation, and social impact - as well as the combination of my filmmaking perspective with Isabelle's experience as an interior architect.
- *Vivaldi Immersive* represents closeness, musical performance, and spatial sound.
- *Faces of Fourni: Bound to the Sea* represents documentary observation, landscape, work, and living environment.
- The portrait of the violin maker represents craft, material, precision, and the concentration of a real working process.
- The work with Takuya represents personality, rhythm, physical energy, and immediate closeness to a musician.

Together, these projects have given me a broad understanding of the medium.

I therefore do not see Apple Immersive Video only as a new camera technology. I see it as a new form of cinematic storytelling that requires different decisions:

- Where does the camera stand?
- How close can it come to a person?
- What meaning does the space have?
- What should the audience be allowed to discover for itself?
- When do I need to guide the viewer, and when should I let go?
- How long can and should a moment last?
- Which story truly becomes stronger through spatial presence?
- Are the people in front of the camera genuinely believable in their behavior?
- Can a scene carry a longer spatial and emotional arc?

My position can therefore be summarized in one sentence:

I create cinematic experiences in which people, spaces, and events are connected in a way that allows the audience to feel not only like a viewer, but like a present part of the situation.

My final conclusion

Apple Immersive Video is especially powerful when five elements come together:

- A person with whom the audience can build a relationship.
- A real action or development.
- A space that has meaning for the story.
- A viewpoint that the audience would not normally be able to experience.
- An authentic presence or acting performance that remains believable across longer spatial sequences.

We Make Belonging fulfills all five conditions.

The film also shows that the value of Apple Immersive Video does not lie only in spectacle. The medium can be equally powerful in:

- art,
- architecture,
- design,
- public space,
- craft,
- documentary observation,
- social issues,
- and quiet human processes.

Another major advantage is that, despite its extraordinary effect, Apple Immersive Video remains film production at its core. I can record real people and real places with a purpose-built camera, work with the material in a familiar post- production environment, and move relatively directly toward a finished experience.

I do not necessarily have to model an artificial world, program a game engine, or develop a complex XR application.

The special sense of presence comes from the combination of spatial accuracy and temporal fidelity: the audience does not only see a believable room, but also experiences natural biological movement within it.

The high spatial and temporal accuracy makes the medium both powerful and unforgiving. It strengthens believable encounters and convincing acting, but it also reveals uncertainty, artificial behavior, and a lack of truthful behavior more quickly.

Casting, trust, rehearsals, and the ability to carry a situation through a longer arc therefore become central parts of immersive design.

This also opens another field beyond storytelling: encountering your own spatial image. People who see themselves in Apple Immersive Video can experience their posture, movement, expression, and impact from an outside perspective that goes far beyond a mirror or conventional video.

This could make the medium relevant in the future for education, coaching, artistic training, and professional self- reflection.

The future will therefore probably not consist only of extreme sports, nature, and large live events. The medium will develop in several directions at the same time:

- premium entertainment,
- live sport and concerts,
- art and culture,
- architecture and spatial documentation,
- portraits and documentary storytelling,
- fictional stories with longer spatial scene arcs,
- corporate and training applications,
- coaching, education, and self-reflection,
- dedicated apps and platforms,
- and supervised experiences in museums, showrooms, and event spaces.

For me, the best description of the medium is therefore not:

"Apple Immersive Video shows the world in extremely high resolution and 3D."

Instead:

With Apple Immersive Video, I can allow people to physically experience a real place, a person, and a moment - and turn a film they have watched into a memory of their own.