Apple Movie Profiles for Spatial and Immersive Media

Profile definitions

Version 0.9 (Beta)

June 9, 2025

Note: The information contained within this document is
         preliminary and is subject to change.

Introduction

This document describes *use* of Apple extensions of, or specialized use of, the QuickTime File Format (a.k.a., QTFF) and the ISO Base Media Format (a.k.a. ISOBMFF) to support spatial and immersive media. Whereas the format details of the syntax and semantics of these extensions are described elsewhere, this document covers restrictions on the use of particular extensions (e.g., limitations on ranges of values, what format tools can be used together). This spec is focused on spatial and immersive media intended to produce a richer experience for the user.

This document introduces the *movie profile* as a collection of allowed media essence carried in tracks and metadata, any associated media signaling, guidance on particular values, and required relationships among the former to ensure suitability for playback to devices or for production for such delivery. Production and delivery are seen as parallel but complementary uses for these format additions. Production may allow more latitude in features used than the same features used for delivery, owing to the more constrained nature of devices served during delivery. The terms *for production* and *for delivery* serve to distinguish these two cases. The term *profile* in the remainder of this document will be understood to be equivalent to *movie profile*, unless prefixed with *media-specific*.

Particular *tools* such as encoded media may require their own media-specific profiles and levels, required metadata or other signaling. The tools themselves (i.e., format features) that are referenced within movie profiles here are enumerated in separate documents. If an implementation supports a particular movie profile's format tools individually, that implementation is well positioned to support a movie profile described here required in the broader ecosystem.

Like the QuickTime File Format (QTFF) upon which it is based, the ISOBMFF format is meant to serve as a container of media using tracks and movie-level structures. The movie format of that media continues to evolve, from the earliest "postage stamp" (i.e., very low-resolution) video, with one- and two-channel uncompressed or barely compressed audio, to modern formats performing remarkable levels of visual compression for 4K and even 8K video, with very rich multichannel, ambisonic and object-based audio encoding. This is coupled with facilities to carry and present captions, such as WebVTT text tracks or closed captions embedded with video. Static and timed metadata can be carried to augment the presented media. Still other kinds of media tracks have been supported and will likely get added.

To support media that delivers rich spatial (or immersive) experiences, the QTFF and ISOBMFF foundations are being extended with new media formats, extensions to supported media formats, and new constructs to inform relationships among the new and earlier supported media. Some of these extensions are specific to their spatial nature, whereas others are fundamental and used by the former. This is all intended to be done in a way—where possible—so an existing ISOBMFF or QTFF player or processor can interact with the spatial media, possibly in a reduced but compatible form, while allowing new playback or processing to take fuller advantage of the newly afforded richness. Some players may require media that conforms to a particular movie profile—or profiles—so those players will confirm the presence of features enumerated in the relevant profile definition described below.

This document references new and updated file format structures to support spatial and immersive media. Some of these structures are accessible through Apple's AVFoundation and CoreMedia framework interfaces, and those serve as the preferred alternative to direct structural access when running on a platform with Apple frameworks available. Those reading or

writing the format directly—pursuant to relevant licensing—should however be able to accomplish their goals with structural descriptions in separate documents.

Another consideration for ISOBMFF is that it is used in a fragmented movie form for HTTP-based delivery technologies such as HTTP Live Streaming. The support in standalone MP4 files and fragmented MP4 resources is much the same, so movie profiles may be applicable to use in streaming, pursuant to signaling in related streaming playlists.

> Note: The words *may*, *should*, and *shall* are used in the conventional specification sense, that is, respectively, to denote permitted, recommended, or required behaviors.

References

[QTFF] QuickTime File Format (QTFF), 2016

[ISOBMFF] ISO/IEC 14496-12:2020 ISO Base Media File Format

[VEXU] ISO Base Media File Format and Apple HEVC Stereo Video—Format Additions, 2025, v.1.9.8

[ISONALU] ISO/IEC 14496-15:2019 "Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format"

[HEVC] ISO/IEC 23008-2:2020 "High efficiency video coding"

[APPLESTEREO] "ISO Base Media File Format and Apple HEVC Stereo Video", 2023, v.0.9

[HEIF] ISO/IEC 23008-12:2022 "Image File Format"

[METADATA] "Video Contour Map Payload Metadata within the QuickTime Movie File Format—Format Additions"

[OMNI] C.Mei and P. Rives. Single View Point Omnidirectional Camera Calibration from Planar Grids. In ICRA, 2007.

Apple Movie Profiles

With format tools described earlier in QuickTime File Format [QTFF], ISOBMF [ISOBMFF], Apple stereo video support [APPLESTEREO] and spatial additions [VEXU] described elsewhere, movie profiles are introduced here in their current version. This document focuses on movie profiles used for spatial and immersive media. This document is not intended to show how features and restrictions on those features might be combined into movie profiles of other use cases. There is typically an overall movie profile, and sometimes further specialization forms for a particular use case (e.g., delivery, production):

- *Apple Projected Media Profile* (shortened as *APMP*), introduced here, enables the comprehensive signaling of video captured with non-rectilinear frame projections. It allows 180°, 360° and wide field-of-view media (e.g., for *action cameras*) to participate in common media workflows, such as capture, playback and editing, using a framework such as AVFoundation. It also provides the basis for system-wide support for these media types across Apple and third-party ecosystems.

- APMP *for delivery* enumerates the allowed format tools for movies that conform to the Apple Projected Media Profile suitable for delivery to conforming devices. This profile enumeration also indicates features that a device or delivery system must support to meet APMP requirements. A device such as a camera or a production process producing movies that conform to this profile are suitable for immediate use by APMP-conforming devices or tools.

- APMP *for production* is related to the delivery profile but allows for some additional format tools that may be useful in creating the APMP-conforming movie for delivery. For example, APMP for production allows for other video codecs than MV-HEVC [HEVC] (e.g., ProRes, non-multiview HEVC) and the use of frame-packed video (e.g., side-by-side, over-under).

- *Spatial Video Profile* includes additional spatial metadata (i.e., baseline, field of view, disparity adjustment, rectilinear projection) [VEXU] alongside a rectilinearly projected stereo video track [APPLEVIDEO] to enable presentation on Apple Vision Pro, with visual treatments that help minimize common causes of stereo-viewing discomfort.

- *Stereoscopic 3D Profile* enables the delivery of 3D rectilinearly projected video with minimal requirements on video format and other signaling. This is typically used for delivery of premium 3D video experiences.

- This document does not try to define a profile for the 2D video use case. If defined, it would cover traditional use of video that is rectilinearly presented with a classic 2D player experience. Such a profile would include anything that is not one of the earlier spatial or immersive movie profiles.

Per-profile feature tables

Each profile section in the remainder of this document gives a high-level view of particular signaling across Apple Movie Profiles documented here. The table in each section focuses on video signaling, but especially the allowed Video Extended Usage [VEXU] signaling for each profile.

Note: In the tables, the presence of a solid circle ('●') indicates that the format feature is required in the profile. A hollow circle ('○') indicates that the format feature is allowed but not required in the profile, and the letter *x* ('x') indicates that the format feature is not allowed in the profile.

2D and 3D, or monoscopic and stereoscopic, respectively

Overview

The profiles described in this document include support for video with one view (*monoscopic*) or two views (*stereoscopic*). Stereoscopic views, presented to the viewer's respective left and right eyes, produce a sense of depth and are sometimes associated with the term *3D*.

In the profiles described below, an illustration of a single view and a stereo pair is shown to convey what the video media carries.

A standalone image that is neither the left nor the right view of a stereo pair and may use a non-rectilinear projection is shown as the following:

An image to be displayed to the user's left eye, perhaps using a non-rectilinear projection (e.g., half equirectangular) is shown like this:


An image to be displayed to the user's right eye, perhaps using a non-rectilinear projection (e.g., half equirectangular) is shown alternatively like this:
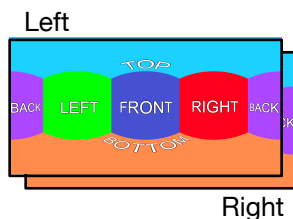

These images provide a legend to diagrams later in this document.
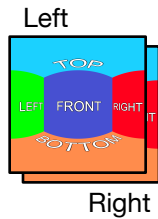
In multiview carriage [APPLESTEREO], the left and right eye views are carried in the same encoded video frame data. This is illustrated with this convention:
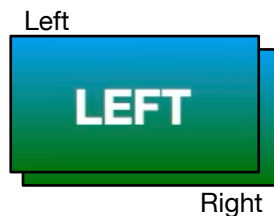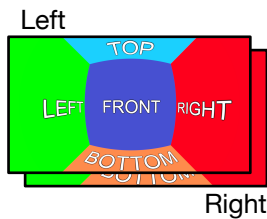

For the full equirectangular (a.k.a., 360°) projection, carriage is indicated in this way for multiview stereo:

Half equirectangular (180°) projection is indicated in this way for multiview stereo:

Left

Right

Wide field-of-view projection is indicated in this way for multiview stereo:

Left

Right

These same images for full equirectangular, half equirectangular and wide field-of-view may be used as single images to indicate one stereo eye's view or the monoscopic view.

3D Movie Profile

Description

The *3D Profile* enables the delivery of 3D rectilinearly projected video with a minimal amount of requirements on video formats and other signaling. This might be used for delivery of premium 3D video experiences or user-generated experiences.

Note: The *Spatial Video Profile* and other profiles add other requirements on the movie's video track signaling, so care should be taken not to prematurely interpret a movie conforming to the 3D Profile as conforming or not conforming to the Spatial Video Profile.

Exemplar

Both the left and right views are encoded in MV-HEVC:

Left

Right
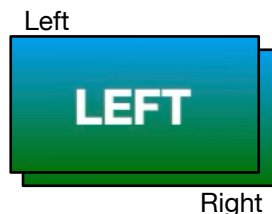
Video requirements

For delivery of 3D video using this profile, the single-video track must use MV-HEVC [HEVC]. Constraints on video resolution and bitrate are not specified.

	Delivery		Production	
	Mono	Stereo	Mono	Stereo
MV-HEVC	Base layer should be visible to HEVC-only player.	Base layer and secondary layer should provide left and right views.	Base layer should be visible to HEVC-only player.	Base layer and secondary layer should provide left and right views.
Non MV-HEVC	x	x	x	Frame-packed video encoded in a non-multiview format (e.g., ProRes, HEVC) can be used.

`VideoExtendedUsageBox` requirements

The following table shows requirements on the structure of the `VideoExtendedUsageBox` for the 3D Movie Profile:

VisualSampleEntry Box type hierarchy$_0$	Stereo–scopic (3D) for delivery	Stereoscopic (3D) for production	Notes
`vexu`$_1$	●	●	
`vexu ▷ must`$_2$	○	○	Version 1 introduces `eyes` and `proj,` so the `must` box does not need to be written.
`vexu ▷ eyes`	●	●	
`vexu ▷ eyes ▷ must`$_2$	○	○	Version 1 introduces `stri` and `hero,` so the `must` box does not need to be written.
`vexu ▷ eyes ▷ stri`	●	●	
`vexu ▷ eyes ▷ hero`	○	○	
`vexu ▷ proj`	●$_4$	●$_4$	If present, the `prji > projection_kind` must be 'rect'. If absent, the projection kind is implied to be `'rect'`.
`vexu ▷ proj ▷ must`$_2$	○	○	Version 1 introduces `prji,` so the `must` box does not need to be written.
`vexu ▷ proj ▷ prji`	● `rect`$_4$	● `rect`$_4$	If specified, the projection_kind must be 'rect'.

VisualSampleEntry Box type hierarchy[0]	Stereo–scopic (3D) for delivery	Stereoscopic (3D) for production	Notes
vexu ▷ pack	x	○	If frame packing is used, box must be present.
vexu ▷ pack ▷ pkin	x	○	If frame packing is used, `view_packing_kind` must be 'side' or 'over', but 'side' is preferred.

Notes: 0 This applies to the ISOBMFF VisualSampleEntry and the QTFF ImageDescription constructs.
1 This box is a VisualSampleEntry (ISOBMFF) or ImageDescription (QTFF) extension.
2 must boxes will be required if a version of this specification introduces a required sibling box, so older readers know they do not understand enough to properly present the containing video.
4 If `proj` is missing, this is equivalent to `proj` being present and `prji` having the `projection_kind` set to 'rect'.

Static metadata requirements

None.

## Timed metadata requirements

None.

Spatial Video Profile

Description

Spatial Video Profile is the native stereo format captured by Apple Vision Pro Camera and the spatial video recording mode of iPhone 15 Pro and 16. These spatial videos can be viewed on Apple Vision Pro. The captured videos include additional spatial metadata (i.e., baseline, field of view, and disparity adjustment) alongside a stereo video track, to enable presentation on Apple Vision Pro.

Only a delivery flavor of Spatial Video Profile is defined. Delivery applies to playback but also to original camera recording.

Exemplar

Both the left and right views are encoded in MV-HEVC:
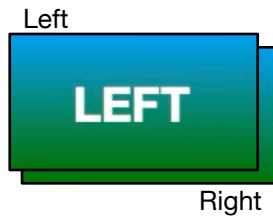
Left

LEFT

Right

Video requirements

For delivery of stereoscopic video using this profile, the single-video track must use MV-HEVC [HEVC]. Constraints on video resolution and bitrate are not specified here.

The Spatial Video Profile requires new signaling in the `VideoExtendedUsageBox` and in the `VisualSampleEntry`. The `VideoExtendedUsageBox` should signal both baseline and disparity adjustment values. The `VisualSampleEntry` [ISOBMFF] (or ImageDescription in [QTFF]) should contain a `HorizontalFieldOfViewBox`.

`VideoExtendedUsageBox` requirements

The following table shows requirements on the structure of the `VideoExtendedUsageBox` for the Spatial Video Profile:

VisualSampleEntry Box type hierarchy[0]	Spatial Video Profile	Notes
vexu[1]	●	
vexu ▷ must[2]	○	
vexu ▷ eyes	●	
vexu ▷ eyes ▷ must[2]	○	
vexu ▷ eyes ▷ stri	●	Both left and right eyes must have views encoded in the MV-HEVC video.
vexu ▷ eyes ▷ hero	○	
vexu ▷ eyes ▷ cams	●	
vexu ▷ eyes ▷ cams ▷ blin	●	The `blin` value should be a positive, non-zero value as it is interpreted as the distance from the optical center of the left-eye camera to the optical center of the right-eye camera.
vexu ▷ eyes ▷ cmfy	●	
vexu ▷ eyes ▷ cmfy ▷ dadj	●	Disparity adjustment must be in the range $[-1, 1]$. A positive adjustment value of `0.02` (2%) is a common default.

VisualSampleEntry Box type hierarchy[0]	Spatial Video Profile	Notes
vexu ▷ proj	● [4]	
vexu ▷ proj ▷ must[2]	○	
vexu ▷ proj ▷ prji	● rect[4]	The projection_kind shall only be 'rect'.

Notes: 0 This applies to the ISOBMFF VisualSampleEntry and the QTFF ImageDescription constructs.
1 This box is a VisualSampleEntry (ISOBMFF) or ImageDescription (QTFF) extension.
2 must boxes will be required if a version of this specification introduces a required sibling box, so that older readers know they do not understand enough to properly present the containing video.
4 If proj is missing, this is equivalent to proj being present and prji having the projection_kind set to 'rect'.

Static metadata requirements

None.

Timed metadata requirements

None.

Apple Projected Media Profile

Apple Projected Media Profile (shortened as *APMP*) enables the comprehensive signaling of video captured with non-rectilinear projections. It allows 180°, 360° and wide field-of-view media to participate in common media workflows, such as capture, playback and editing, using a framework such as AVFoundation. It also provides the basis for system-wide support for these media types across Apple and third-party ecosystems.

It is further specialized for two use cases: delivery and production. The two subordinate use cases are described in the next two sections.

Apple Projected Media Profile for delivery

Description

APMP *for delivery* enumerates the allowed format tools for movies that conform to the Apple Projected Media Profile suitable for delivery to conforming devices. This profile enumeration also indicates features a device or delivery system must support to meet APMP requirements. A device such as a camera or a production process that produces APMP-conforming movies and the delivery use case are suitable for immediate use by APMP-conforming devices or tools.

While targeting delivery to an APMP-conforming device, a device can capture APMP for delivery movies.

Exemplar

For monoscopic, the 2D view can be encoded to HEVC with a full equirectangular, half equirectangular, or parametric projection:

or   or   or
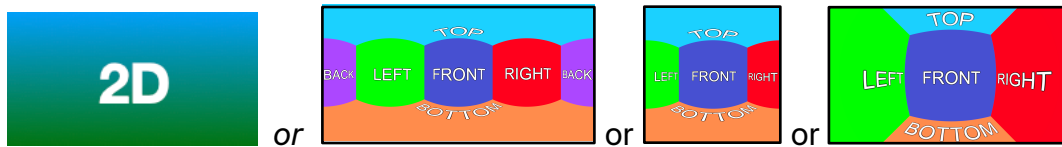
For stereo, both the left and right views can be encoded in MV-HEVC with a full equirectangular, half equirectangular or parametric projection:

or   *or*   *or*
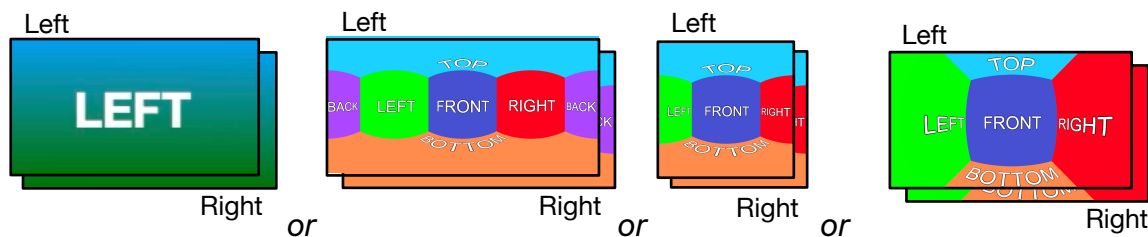
Video requirements

For delivery of stereoscopic video using this profile, a single video track must use MV-HEVC [HEVC]. Constraints on video resolution and bitrate are not specified here.

`VideoExtendedUsageBox` requirements

The following table shows requirements on the structure of the `VideoExtendedUsageBox` for the Apple Project Media Profile for Delivery:

VisualSampleEntry Box type hierarchy$_0$	APMP for delivery (stereoscopic)	APMP for delivery (monoscopic)	Notes
Video coding	MV−HEVC		
vexu$_1$	●	●	
vexu ▷ must$_2$	○	○	
vexu ▷ eyes	●	x	
vexu ▷ eyes ▷ must$_2$	○	x	

12

VisualSampleEntry Box type hierarchy$_0$	APMP for delivery (stereoscopic)	APMP for delivery (monoscopic)	Notes
vexu ▷ eyes ▷ stri	○	x	
vexu ▷ eyes ▷ hero	○	x	
vexu ▷ eyes ▷ cams	○	x	
vexu ▷ eyes ▷ cams ▷ blin	○	x	This is optional in APMP for delivery.
vexu ▷ eyes ▷ cmfy	x	x	
vexu ▷ eyes ▷ cmfy ▷ dadj	x	x	This is not supported in APMP for delivery.
vexu ▷ proj	●	●	
vexu ▷ proj ▷ must$_2$	○	○	
vexu ▷ proj ▷ prji	● equi| hequ| prim	● equi| hequ| prim	Projections can be 360° (equi), 180° (hequ) or parametric immersive (prim).
vexu ▷ pack	x	x	Delivery does not support frame packing.
vexu ▷ pack ▷ must$_2$	x	x	
vexu ▷ pack ▷ pkin	x (only MV-HEVC)	x	
vexu ▷ lnsc$_3$	○	○	Only parametric immersive projection should characterize the stereo lens pair.
vexu ▷ lnsc ▷ lens	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd	○	○	
vexu ▷ lnsc ▷ lens ▷ rdim	○	○	
vexu ▷ lnsc ▷ lens ▷ lnin	○	○	
vexu ▷ lnsc ▷ lens ▷ ldst	○	○	
vexu ▷ lnsc ▷ lens ▷ lnex	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ must$_2$	○	○	

VisualSampleEntry Box type hierarchy[0]	APMP for delivery (stereoscopic)	APMP for delivery (monoscopic)	Notes
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ corg	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ cxfm	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ cxfm ▷ uqua	○	○	
hfov[1]	○	○	

Notes: 0 This applies to the ISOBMFF VisualSampleEntry and the QTFF ImageDescription constructs.
1 This box is a VisualSampleEntry (ISOBMFF) or ImageDescription (QTFF) extension.
2 must boxes will be required if a version of this specification introduces a required sibling box, so that older readers know they do not understand enough to properly present the containing video.
3 The lens collection (lnsc) is required for the ParametricImmersive projection (`prim`) but is not required for other projection kinds.

Static metadata requirements

None.

Timed metadata requirements

None.

# Apple Projected Media Profile for production

Description

APMP *for production* is closely related to APMP for delivery. The production focus allows for additional flexibility in format tools that may be useful in creating the movie that will be used to produce content conforming to APMP for delivery. For example, APMP for production allows for video codecs other than MV-HEVC [HEVC] and the use of frame-packed video (side-by-side, over–under) with non-layered codecs. This flexibility means that a "2D" codec such as HEVC (without use of MV-HEVC extensions) or Apple ProRes is suitable. Other 2D codecs can be used.

Exemplar

During the production process, it is possible to produce movies that are not yet converted for delivery. Consequently, there is no requirement that the video track use MV-HEVC. It can use other production video codecs such as ProRes, a RAW video format or HEVC without multiview extensions.

For monoscopic production, the video track can use any available video codec throughout the production workflow.
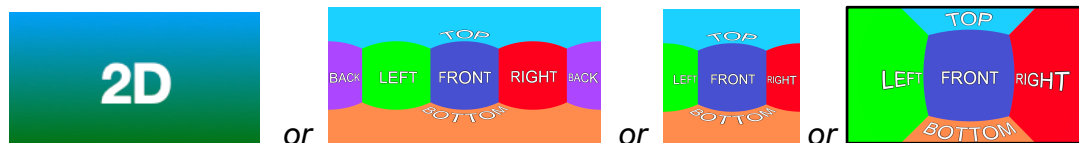
For stereo production, both the left and right views may be encoded in MV-HEVC for rectangular, full equirectangular, half equirectangular and wide field-of-view:
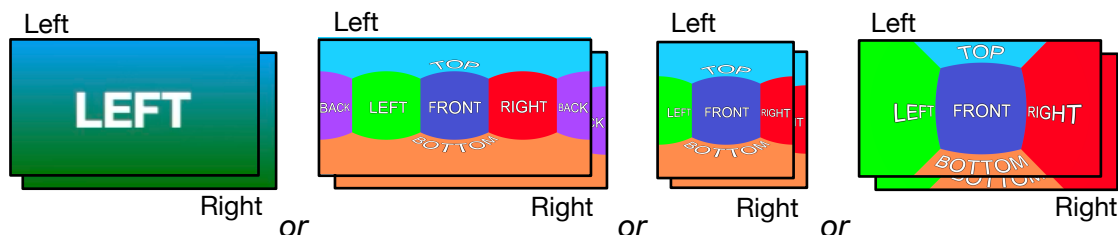
If a non-multiview video codec is used, the use of a frame-packed layout in the encoded video frame is allowed. In this case, a more traditional 2D codec such as ProRes, a RAW format or HEVC without multiview extensions can be used. To signal the stereo, side-by-side frame packing is allowed:
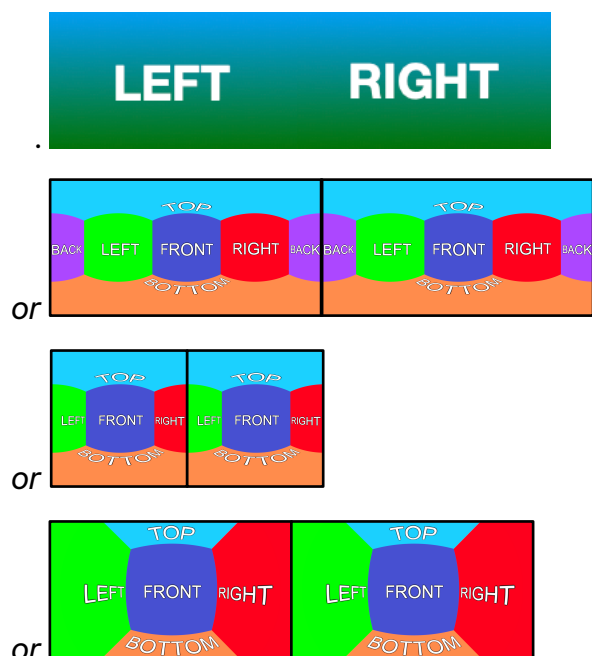
(Note: The black frame around each view is only to indicate separate views. It is not an indication of framing in the actual image pixels.)

Over-under frame packing layout is also allowed:
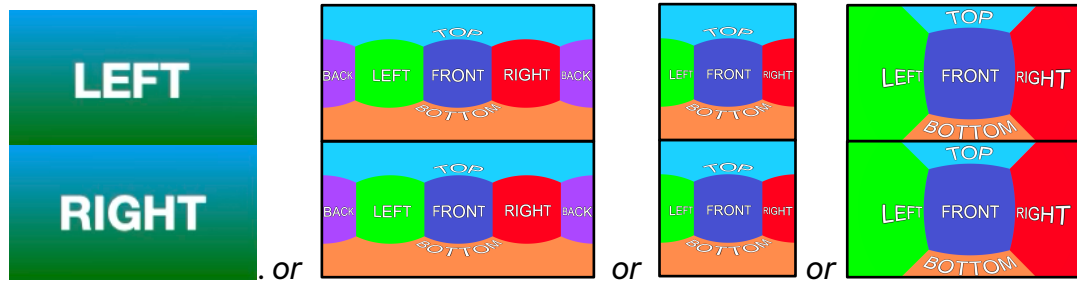
. *or*   *or*   *or*

(Note: The black frame around each view is only to indicate separate views. It is not an indication of framing in the actual image pixels.)

Video requirements

`VideoExtendedUsageBox` requirements

The following table shows requirements on the structure of the `VideoExtendedUsageBox` for the Apple Project Media Profile for production:

VisualSampleEntry Box type hierarchy$_0$	APMP for production (stereoscopic)	APMP for production (monoscopic)	Notes
Video coding	MV-HEVC Other codecs (e.g., ProRes, a RAW format, HEVC) with additional frame packing signaling.	Not MV-HEVC (e.g., ProRes, a RAW format, HEVC).	Multiview and 2D codecs may both be used for stereoscopic video. If a 2D codec is used, frame packing should be used. Otherwise, the multiview codec carries the left- and right-eye views. 2D codecs should be used for monoscopic video. There are no restrictions on the specific 2D video codecs used, but it is recommended that it be one of Apple ProRes, a compatible RAW encoding, or HEVC (without multiview extensions being used). Because this is for production, there is no limit on the bit rate allowed in a production workflow.

VisualSampleEntry Box type hierarchy[0]	APMP for production (stereoscopic)	APMP for production (monoscopic)	Notes
`vexu`[1]	●	●	
`vexu ▷ must`[2]	○	○	
`vexu ▷ eyes`	○	x	
`vexu ▷ eyes ▷ must`[2]	○	x	
`vexu ▷ eyes ▷ stri`	○	x	The `eyes ▷ stri` box can be used to signal that neither the left nor the right eye is used, but it may be better not to signal anything stereo-related.
`vexu ▷ eyes ▷ hero`	○	x	
`vexu ▷ eyes ▷ cams`	○	x	
`vexu ▷ eyes ▷ cams ▷ blin`	○	x	
`vexu ▷ eyes ▷ cmfy`	x	x	
`vexu ▷ eyes ▷ cmfy ▷ dadj`	x	x	
`vexu ▷ proj`	●	●	
`vexu ▷ proj ▷ must`[2]	○	○	
`vexu ▷ proj ▷ prji`	● `equi|hequ|prim`	● `equi|hequ|prim`	
`vexu ▷ pack`	○	x	Only frame-packed stereo video requires the `pack` box hierarchy.
`vexu ▷ pack ▷ must`[2]	○	x	
`vexu ▷ pack ▷ pkin`	○ `view_packing_kind` can also be `side|over`	x	Only frame-packed stereo video requires the pack box hierarchy. The `view_packing_kind` can be 'side' or 'over'. 'side' is recommended.
`vexu ▷ lnsc`[5]	○	○	Only parametric immersive projection should characterize the stereo lens pair.

VisualSampleEntry Box type hierarchy[0]	APMP for production (stereoscopic)	APMP for production (monoscopic)	Notes
vexu ▷ lnsc ▷ lens	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd	○	○	
vexu ▷ lnsc ▷ lens ▷ rdim	○	○	
vexu ▷ lnsc ▷ lens ▷ lnin	○	○	
vexu ▷ lnsc ▷ lens ▷ ldst	○	○	
vexu ▷ lnsc ▷ lens ▷ lnex	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ must[2]	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ corg	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ cxfm	○	○	
vexu ▷ lnsc ▷ lens ▷ lnhd ▷ cxfm ▷ uqua	○	○	
hfov[1]	○	○	The horizontal field of view may be specified.

Notes: 0 This applies to the ISOBMFF VisualSampleEntry and the QTFF ImageDescription constructs.
1 This box is a VisualSampleEntry (ISOBMFF) or ImageDescription (QTFF) extension.
2 must boxes will be required if a version of this specification introduces a required sibling box, so that older readers know they do not understand enough to properly present the containing video.
4 If proj is missing, this is equivalent to `proj` being present and `prji` having the `projection_kind` set to `'rect'`.
5 The lens collection (lnsc) is required for the ParametricImmersive projection (`prim`) but is not required for other projection kinds.

Static metadata requirements

None.

Timed metadata requirements

None.

Conclusion

This document introduces movie profiles describing the use of extensions to the QuickTime (.mov) and ISOBMFF movie formats for particular use cases. The particular format tools allowed by these movie profiles are described outside this document. This document is meant to enumerate the tools used and any limits on particular tools. Among the movie profiles described are those already in use for Apple Vision Pro content delivery and a new Apple Projected Media Profile.

Document Revision History

Date	Revision	Notes
2025-06-09	0.9	First version