**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 N17372**

**January 2018, Gwangju, KR**

|  |  |
| --- | --- |
| **Source:** | Requirements |
| **Title** | How OMAF fulfils MPEG-I phase 1a requirements |
| **Editors** | Sachin Deshpande, Ye-Kui Wang, Thomas Stockhammer |
| **Public** | Yes |

# Introduction

This document provides a summary of MPEG-I Phase 1A requirements fulfilment by OMAF 23090-2 FDIS.

# Requirements Fulfilment

The table below lists numbered requirements from MPEG-I Phase 1A as documented in N16773 [1] in the left column and the clause number from OMAF 23090-2 [2] which fulfils the requirement in the right column.

**Table 2.1 – MPEG-I Phase 1 A Requirements fulfilment by OMAF 23090-2**

|  |  |
| --- | --- |
| **MPEG-I Phase 1 A Requirement # per N16773** | **OMAF 23090-2 FDIS clause** |
| 1. The Specification shall provide for interoperable exchange of VR360 content. | All normative clauses of 23090-2, in particular clauses 5-11. |
| 1. The Specification shall avoid providing multiple tools for the same functionality to reduce implementation burden and improve interoperability. | Clause 5-11 |
| 1. The Specification shall enable good quality and performance. | Clause 5-11 |
| 1. The Specification shall enable full interoperability between services/content and clients. | Interoperability is achieved via media profiles (Clause 10) and presentation profiles (clause 11). |
| * 1. The Specification shall contain a very low number of fully specified interoperability points that include what is traditionally known as Profile and Level information. | Two presentation profiles: OMAF viewport-independent presentation profile (Clause 11.1), OMAF viewport-dependent profile (Clause 11.2)  Three video profiles:  HEVC-based viewport-independent OMAF video profile (Clause 10.1.2)  HEVC-based viewport-dependent OMAF video profile (Clause 10.1.3)  AVC-based viewport-dependent OMAF video profile (Clause 10.1.4)  Two audio profiles:  OMAF 3D audio baseline profile (Clause 10.2.2)  OMAF 2D audio legacy profile (Clause 10.2.3)  Two image profiles:  OMAF HEVC image profile (Clause 10.3.2, 10.3.3)  OMAF legacy image profile (Clause 10.3.2, 10.3.4)  Two Timed text profiles:  OMAF IMSC1 timed text profile (Clause 10.4.2)  OMAF WebVTT timed text profile (Clause 10.4.3) |
| * + 1. The existence of more than one interoperability point shall be justified if intended to target devices with different capabilities. | Media profiles are mainly targeted to legacy devices and current state-of-the-art devices with different capabilities (Clause 10)  Presentation profiles are mainly targeted for viewport independent and viewport dependent presentation (Clause 11). |
| * 1. Interoperability points shall address a Media Profile including: | Media profiles (Clause 10) |
| * + 1. file format tracks and elementary stream | Media profiles (Clause 10)  File format track constraints: Clause 10.1.2.4, 10.1.3.3, 10.1.4.3, 10.2.2.3, 10.2.3.3  Elementary stream constraints:  Clause 10.1.2.2, 10.1.3.2, 10.1.4.2, 10.2.2.2, 10.2.3.2 |
| * + 1. rendering: The Specification shall provide interoperability points that include equirectangular projection. Other projection formats shall only be included if there are proven benefits and industry support; | HEVC-based viewport-independent OMAF video profile (Clause 10.1.2) includes equirectangular projection  HEVC-based viewport-dependent OMAF video profile (Clause 10.1.3) includes equirectangular projection and cubemap projection. |
| * 1. Interoperability points shall address a Presentation Profile for a full VR experience including different media (Video, Audio and Subtitles), enabling their temporal synchronization and spatial alignment | Two presentation profiles: OMAF viewport-independent presentation profile (Clause 11.1), OMAF viewport-dependent profile (Clause 11.2) |
| * 1. These interoperability points shall enable conformance to be tested, inside and outside of MPEG. | Clause 11, Clause 10 allow conformance definition.  External SDOs (e.g. VR-IF) have defined guidelines based on 23090-2 profiles. |
| * 1. The Specification may contain partial interoperability points (e.g., a file format box, a visual media profile) at a lower level of granularity. | Clauses 6, 7, 8 provide features including file format boxes, DASH descriptors, MMT messages |
| * 1. The Specification may contain optional elements (like a description of the Director’s recommended viewport) when such options do not affect basic interoperability; Profiles can make such features mandatory but these features are not necessarily included in a Profile. | Media Profiles in Clause 10, Presentation Profiles in Clause 11, Recommended viewport in Clause 7.5.5 |
| * 1. The specification shall define at least one media profile for audio. | Two audio profiles:  OMAF 3D audio baseline profile (Clause 10.2.2)  OMAF 2D audio legacy profile (Clause 10.2.3) |
| * 1. The specification shall define at least one media profile for video. | Three video profiles:  HEVC-based viewport-independent OMAF video profile (Clause 10.1.2)  HEVC-based viewport-dependent OMAF video profile (Clause 10.1.3)  AVC-based viewport-dependent OMAF video profile (Clause 10.1.4) |
| * 1. The specification shall define at least one presentation profile that includes one audio and one video media profile. | Two presentation profiles each including one audio and one video media profile: OMAF viewport-independent presentation profile (Clause 11.1), OMAF viewport-dependent profile (Clause 11.2) |
| 1. The Specification should take into account the capabilities of high quality devices such as HMDs that are on the market today (including Vive, Oculus, Gear VR, and Daydream) or that are on the market by the time the specification is stable, i.e., Q4 2017. | Clause 5-11 |
| 1. The Specification shall support the representation, storage, delivery and rendering of: | Representation – Clause 5, and 6  Storage – Clause 7  Delivery – Clause 8, 9 |
| * 1. Omnidirectional (up to 360° spherical) coded image/video (monoscopic and stereoscopic) with 3 DoF | Clause 5-11 |
| * 1. Both 3D and 2D audio | Audio Profiles Clause 10.2  Clause 10.2.2: 3D audio  Clause 10.2.3: 2D audio |
| 1. The specification shall work with existing MPEG storage and delivery formats | Clause 7 works with ISOBMFF storage,  Clause 8 works with DASH,  Clause 9 works with MMT |
| 1. The Specification shall support temporal synchronization and spatial alignment between different media types, in particular between audio and video. | Supported via use of ISOBMFF storage (Clause 7), DASH (Clause 8) and MMT delivery (Clause 9) |
| 1. The Specification shall support metadata for describing initial viewpoints and for the playback of omnidirectional video/image and audio according to that metadata. | Initial viewing orientation signalling for ISOBMFF in clause 7.5.4,  Initial viewing orientation item property signalling for omnidirectional images in clause 7.7.9,  Carriage of timed metadata in DASH in clause 8.2.3,  Initial viewing orientation sample signalling in MMT VR information asset descriptor in Table 9.2 |
| 1. The Specification shall support the following interfaces: |  |
| * 1. encoding and decoding for each media type | Media Profiles Clause 10 defines supported codecs (encoding, decoding). |
| * 1. delivery for download and streaming. | Clause 8 for DASH delivery, Clause 9 for MMT delivery |
| 1. The Specification shall enable applications to use hardware-supported or pre-installed independently manufactured decoders and renderers through defined MPEG conformance points. | Media Profiles Clause 10 define supported codecs (encoding, decoding).  Media profiles (Clause 10) and Presentation profiles (Clause 11) define conformance points. |
| 1. The Specification shall support viewport-dependent processing (this may include delivery, decoding and rendering). | OMAF viewport-dependent profile (Clause 11.2) |
| * 1. The Specification shall support dynamically changing viewports. | Recommended viewport timed metadata track (Clause 7.5.5) |
| * 1. The Specification should enable responsiveness to changing viewport in a way that doesn’t detract from the immersive experience. | Recommended viewport timed metadata track (Clause 7.5.5) |
| 1. The Specification shall support at least one Presentation Profile that requires support for neither viewport-dependent delivery nor viewport-dependent decoding. | Clause 11.1 OMAF viewport-independent baseline presentation profile |
| Note: it is obvious that there will be viewport-dependent rendering, both for visual and audio components |  |
| Delivery |  |
| 1. The Specification shall support the following methods of distribution: |  |
| * 1. File-based delivery | Clause 7 |
| * 1. DASH-based streaming | Clause 8 |
| * 1. MMT-based streaming | Clause 9 |
| Visual |  |
| 1. The Specification shall enable content exchange with high visual perceptual quality. | HEVC Main 10 Level 5.1 supported by HEVC-based viewport-independent OMAF video profile (Clause 10.1.2), HEVC-based viewport-dependent OMAF video profile (Clause 10.1.3) is capable of providing high visual perceptual quality |
| * 1. Taking the display resolution of existing headsets into consideration, the Specification shall support a visible viewport resolution beyond which the increase in resolution is no longer noticeable on these headsets.   Note: This may equate to a source resolution (for the full 360 video) of around 6k x 3k or 8k x 4k for equirectangular pictures (where the viewport is only the visible part of the panorama at a given point of time). | HEVC Main 10 Level 5.1 supported by HEVC-based viewport-independent OMAF video profile (Clause 10.1.2), HEVC-based viewport-dependent OMAF video profile (Clause 10.1.3) provides adequate resolution to existing headsets (e.g. Oculus Rift, HTC Vive, GearVR, etc.) |
| * 1. The Specification shall support a framerate of at least 60fps | HEVC Main 10 Level 5.1 Profile Level supported by HEVC-based viewport-independent OMAF video profile (Clause 10.1.2), HEVC-based viewport-dependent OMAF video profile (Clause 10.1.3) supports framerate higher than 60 fps. |
| 1. The Specification shall support distribution of full panorama resolutions beyond 4K (e.g. 8K, 12K), to decoders capable of decoding only up to 4K@60fps, if sufficient interoperability can be achieved. | HEVC Main 10 Level 5.1 used by media profiles (Clause 10.1.2, 10.1.3) supports resolutions beyond 4K.  Annex D Clause D.6 (informative) describes achieving “5K”, “6K” effective equirectangular and cubemap resolution using viewport dependent video media profiles (Clause 10.1.3, 10.1.4) |
| 1. The Specification shall support metadata for the rendering of spherical video on a 2D screen. | Omnidirectional metadata signalling in Clause 7. |
| 1. The Specification shall support fisheye-based video with a configuration of 2 cameras. | Clause 6 |
| 1. The Specification shall support encoding of equirectangular projection (ERP) maps for monoscopic and stereoscopic video, in an efficient manner. | Clause 5.2.1, 5.2.2 |
| * 1. Other projection maps than ERP for distribution should only be provided if consistent benefits over ERP is demonstrated. | Clause 5.2.1, 5.2.3 additionally supports cubemap projection.  Table 5.1 provides future extensibility for omnidirectional projection formats. |
| Audio |  |
| 1. Each audio media profile in the Specification shall: |  |
| * 1. support immersive rendering with sufficiently low latency   Note: this is related to requirement 12.2 | Clause 10.2.2 -OMAF 3D Audio baseline profile which uses MPEG-H 3D audio.  Latency aspects also briefly described in Clause 10.2.2.1 |
| * 1. support Excellent sound quality (as assessed per ITU-R BS.1534) | Clause 10.2.2 -OMAF 3D Audio baseline profile which uses MPEG-H 3D audio. |
| * 1. support binauralization   Note: binauralization implies adaptivity to user head motion, such that the user experiences directional audio that is consistent with such head motion. | Clause 10.2.2 -OMAF 3D Audio baseline profile which uses MPEG-H 3D audio.  Binauralization aspect also briefly described in Clause 10.2.2.1 |
| 1. There may be one audio media profile that supports only 2D audio to cater to existing devices. | OMAF 2D audio legacy profile (Clause 10.2.3) |
| 1. All other audio media profiles defined in the Specification shall: | Clause 10.2.2 -OMAF 3D Audio baseline profile which uses MPEG-H 3D audio. |
| * 1. support 3D Audio distribution, decoding & rendering. |
| * 1. support immersive content, e.g. 12ch or 3rd order Ambisonics, |
| * 1. support a combination of diegetic and non-diegetic content sources. |
| * 1. be capable to ingest and carry all content types: |
| * + 1. audio channels, |
| * + 1. audio objects, |
| * + 1. scene-based audio, |
| * + 1. and combinations of the above. |
| * 1. be able to carry dynamic meta-data for combining, presenting and rendering all content types. |
|  |  |
| Security |  |
| 1. The Specification shall not preclude: | Clause 5-11 |
| * 1. Decoding and rendering to support secure media pipelines | Clause 5-11 |
| * 1. Efficient distribution for multiple DRM systems (e.g. using common encryption) | DASH (Clause 8) / MMT (Clause 9) delivery enables a secure media pipeline. |
| 1. The Specification should enable a secure media pipeline to be implemented. | ISOBMFF Storage (Clause 7), DASH (Clause 8) / MMT (Clause 9) delivery enables a secure media pipeline.  CMAF integration of media profiles (Annex C)- Clause 7.1.2 supports ‘cbcs’ scheme. |

# Conclusion

It can be observed by the table 2.1 that all the requirements for MPEG-I Phase 1A are fulfilled by OMAF 23090-2.

# References

1. MPEG, N16773, “Requirements for OmniDirectional MediA Format,” April 2017
2. MPEG, N17235, ISO/IEC 23090-2, “Information technology — Coded representation of immersive media (MPEG-I) — Part 2: Omnidirectional media format,” December 2017