

---

---

**Information technology — Coded  
representation of immersive media —  
Part 2:  
Omnidirectional media format**

*Technologies de l'information — Représentation codée de média  
immersifs —*

*Partie 2: Format de média omnidirectionnel*





**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2021

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
CP 401 • Ch. de Blandonnet 8  
CH-1214 Vernier; Geneva  
Phone: +41 22 749 01 11  
Email: [copyright@iso.org](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

Foreword.....	vii
Introduction.....	viii
<b>1 Scope .....</b>	<b>1</b>
<b>2 Normative references .....</b>	<b>1</b>
<b>3 Terms, definitions, abbreviated terms and conventions .....</b>	<b>2</b>
3.1 Terms and definitions .....	2
3.2 Abbreviated terms .....	10
3.3 Conventions .....	10
3.3.1 Arithmetic operators and mathematical functions.....	10
3.3.2 Order of operation precedence.....	12
3.3.3 Range notation .....	12
3.3.4 Variables .....	13
3.3.5 Processes .....	13
3.3.6 Conventions for indicating the number of boxes in tables.....	13
<b>4 Overview.....</b>	<b>13</b>
4.1 Organization of this clause.....	13
4.2 Overall architecture .....	14
4.3 Projected omnidirectional video/images.....	16
4.3.1 General.....	16
4.3.2 Stitching, rotation, projection, and region-wise packing .....	16
4.4 Fisheye omnidirectional video/images.....	17
4.5 Mesh omnidirectional video .....	18
4.6 Streaming methods for omnidirectional video .....	18
4.6.1 Overview.....	18
4.6.2 Tile-based streaming with viewport-specific author-driven binding .....	20
4.6.3 Tile-based streaming with free-viewport author-driven binding.....	20
4.6.4 Tile-based streaming with late binding .....	21
4.7 Additional functionalities .....	22
4.8 Conformance and interoperability .....	23
4.8.1 General.....	23
4.8.2 Media profiles.....	23
4.8.3 Presentation profiles .....	25
4.8.4 Toolset brands.....	25
4.8.5 Summary of referenceable code points .....	26
<b>5 Omnidirectional video projection and region-wise packing .....</b>	<b>32</b>
5.1 Coordinate system .....	32
5.2 Omnidirectional projection formats .....	33
5.2.1 General.....	33
5.2.2 Equirectangular projection for one sample location .....	34
5.2.3 Cubemap projection for one sample location.....	34
5.3 Conversion from the local coordinate axes to the global coordinate axes .....	36
5.4 Region-wise packing formats .....	37
5.4.1 General.....	37
5.4.2 Conversion of one sample location for rectangular region-wise packing .....	38
<b>6 Fisheye omnidirectional video.....</b>	<b>39</b>
6.1 General.....	39
6.2 The <code>FisheyeVideoEssentialInfoStruct()</code> syntax structure.....	39
6.2.1 Syntax.....	39
6.2.2 Semantics.....	40
6.3 The <code>FisheyeVideoSupplementalInfoStruct()</code> syntax structure.....	43
6.3.1 Syntax.....	43
6.3.2 Semantics.....	44

<b>7</b>	<b>Omnidirectional media storage and metadata signalling in the ISOBMFF</b>	<b>48</b>
7.1	Generic extensions to the ISOBMFF	48
7.1.1	Indication of a track not intended to be presented alone	49
7.1.2	Association of timed metadata tracks with media tracks or track groups	49
7.1.3	Clarifications on the stereo video box	49
7.1.4	Generic sub-picture track grouping extensions	49
7.1.5	Track reference indicating a track providing shadow sync samples	53
7.1.6	Media offset box	54
7.2	Generic extensions to ISO/IEC 14496-15	55
7.2.1	Containing of SpatialRelationship2DDescriptionBox for HEVC tile base track and HEVC tile tracks	56
7.3	OMAF-specific extensions to the ISOBMFF	56
7.3.1	Sync samples in timed metadata tracks	56
7.4	OMAF-specific extensions to ISO/IEC 14496-15	56
7.4.1	Coverage information box in an HEVC tile base track	56
7.5	Structures and semantics that are common for video tracks and image items	56
7.5.1	Semantics of sample locations within a decoded picture	56
7.5.2	Projection format structure	60
7.5.3	Region-wise packing structure	60
7.5.4	Rotation structure	68
7.5.5	Content coverage structure	69
7.5.6	Sphere region structure	70
7.6	Restricted video schemes for omnidirectional video	73
7.6.1	Scheme types	73
7.6.2	Projected omnidirectional video box	78
7.6.3	Fisheye omnidirectional video box	79
7.6.4	Region-wise packing box	80
7.6.5	Rotation box	80
7.6.6	Coverage information box	81
7.6.7	Mesh omnidirectional video box	81
7.6.8	Mesh box	82
7.7	Timed metadata for sphere regions	84
7.7.1	General	84
7.7.2	Sample entry	85
7.7.3	Sample format	86
7.7.4	Initial viewing orientation	86
7.7.5	Recommended viewport	88
7.7.6	Timed text sphere location metadata	91
7.8	Signalling of region-wise quality ranking	92
7.8.1	General	92
7.8.2	Spherical region-wise quality ranking	92
7.8.3	2D region-wise quality ranking	94
7.9	Storage of omnidirectional images	96
7.9.1	General	96
7.9.2	Frame packing item property	96
7.9.3	Projection format item property	97
7.9.4	Essential fisheye image item property	98
7.9.5	Supplemental fisheye image item property	99
7.9.6	Region-wise packing item property	99
7.9.7	Rotation item property	100
7.9.8	Coverage information item property	100
7.9.9	Initial viewing orientation item property	101
7.10	Storage of timed text for omnidirectional video	102
7.10.1	General	102
7.10.2	OMAF timed text configuration box	102
7.10.3	IMSC1 tracks	104
7.10.4	WebVTT tracks	105
7.11	ERP region timed metadata	105
7.11.1	General	105
7.11.2	Sample entry format	106

7.11.3	Semantics.....	106
7.11.4	Sample format .....	107
7.11.5	Generating ERP region metadata.....	108
7.12	Storage and signalling of viewpoints for omnidirectional video and images .....	108
7.12.1	Viewpoint information structures.....	108
7.12.2	Viewpoint entity grouping.....	117
7.12.3	Timed metadata for viewpoints .....	119
7.13	Storage of omnidirectional video in sub-picture tracks.....	123
7.13.1	General.....	123
7.13.2	Projected omnidirectional video.....	124
7.13.3	Indication of composition pictures being packed pictures or projected pictures .....	125
7.13.4	Fisheye omnidirectional video.....	125
7.14	Storage and signalling of overlays for omnidirectional video and images.....	125
7.14.1	General.....	125
7.14.2	Overlay structure .....	128
7.14.3	Overlay control structures.....	129
7.14.4	Overlay configuration box.....	139
7.14.5	Overlay item property .....	140
7.14.6	Overlay timed metadata track.....	140
7.14.7	Entity groups.....	142
7.14.8	Overlay alpha auxiliary image.....	144
7.15	Signalling of viewing space information.....	145
7.15.1	General.....	145
7.15.2	Viewing space structure.....	145
7.15.3	Viewing space box .....	148
7.15.4	Viewing space item property.....	148
7.15.5	Time varying immersive viewing space signalling.....	148
7.16	Mapping of rectangular regions to the 3D mesh.....	149
7.16.1	General.....	149
7.16.2	Tile mesh sample grouping .....	149
7.16.3	Rectangular region structure.....	151
7.16.4	Projection of a sample location onto the 3D mesh.....	152
<b>8</b>	<b>Omnidirectional media encapsulation and signalling in DASH .....</b>	<b>153</b>
8.1	Architecture of DASH delivery in OMAF .....	153
8.2	Usage of DASH in OMAF .....	155
8.2.1	General.....	155
8.2.2	Signalling of stereoscopic frame packing.....	155
8.2.3	Carriage of timed metadata.....	155
8.2.4	Associating Adaptation Sets or Representations with each other .....	156
8.3	DASH MPD descriptors for omnidirectional media in the namespace "urn:mpeg:mpegI:omaf:2017".....	157
8.3.1	XML namespace and schema .....	157
8.3.2	Signalling of projection type information .....	157
8.3.3	Signalling of region-wise packing type.....	158
8.3.4	Signalling of content coverage .....	159
8.3.5	Signalling of spherical region-wise quality ranking.....	162
8.3.6	Signalling of 2D region-wise quality ranking .....	168
8.3.7	Signalling of fisheye omnidirectional video.....	173
8.4	Carriage of images .....	173
8.4.1	General.....	173
8.4.2	Format and constraints for Segments.....	174
8.5	DASH MPD descriptors for omnidirectional media in the namespace "urn:mpeg:mpegI:omaf:2020".....	174
8.5.1	XML namespace and schema .....	174
8.5.2	Signalling of association.....	174
8.5.3	Signalling of viewpoints .....	176
8.5.4	Signalling of sub-picture composition identifier and its attributes.....	183
8.5.5	Signalling of overlays .....	184
8.5.6	Entity to group descriptor .....	186

8.6	Segment formats .....	188
8.6.1	Initialization Segment for OMAF base track .....	188
8.6.2	Tile Index Segment.....	192
8.6.3	Tile Data Segment.....	193
<b>9</b>	<b>Omnidirectional media encapsulation and signalling in MMT .....</b>	<b>194</b>
9.1	Architecture of MMT delivery in OMAF .....	194
9.2	OMAF signalling in MPEG composition information.....	195
9.3	VR application-specific MMT signalling.....	195
9.3.1	General .....	195
9.3.2	MMT signalling .....	196
<b>10</b>	<b>Media profiles .....</b>	<b>212</b>
10.1	Video profiles.....	212
10.2	Audio profiles .....	243
10.3	Image profiles.....	252
10.4	Timed text profiles .....	257
<b>11</b>	<b>Presentation profiles .....</b>	<b>258</b>
11.1	OMAF viewport-independent baseline presentation profile .....	258
11.1.1	General .....	258
11.1.2	ISO Base Media File Format constraints.....	259
11.2	OMAF viewport-dependent baseline presentation profile .....	259
11.2.1	General .....	259
11.2.2	ISO Base Media File Format constraints.....	259
<b>12</b>	<b>OMAF toolset brands.....</b>	<b>260</b>
12.1	Overlay toolset brand.....	260
12.1.1	Overview.....	260
12.1.2	ISO Base Media File Format constraints.....	260
12.1.3	OMAF player operation.....	260
12.2	Viewpoint toolset brand .....	260
12.2.1	Overview.....	260
12.2.2	ISO Base Media File Format constraints.....	260
12.2.3	OMAF player operation.....	260
12.3	Non-linear storyline toolset brand.....	260
12.3.1	Overview.....	260
12.3.2	ISO Base Media File Format constraints.....	261
12.3.3	OMAF player operation.....	261
	<b>Annex A (normative) OMAF DASH schema.....</b>	<b>262</b>
	<b>Annex B (normative) DASH integration of media profiles .....</b>	<b>266</b>
	<b>Annex C (normative) CMAF integration of media profiles .....</b>	<b>279</b>
	<b>Annex D (informative) Viewport-dependent omnidirectional video processing.....</b>	<b>282</b>
	<b>Annex E (informative) DASH MPD examples.....</b>	<b>316</b>
	<b>Annex F (informative) MMT signalling examples.....</b>	<b>320</b>
	<b>Annex G (normative) Expected behaviour of OMAF player .....</b>	<b>322</b>
	<b>Bibliography .....</b>	<b>332</b>

## Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see [www.iso.org/directives](http://www.iso.org/directives) or [www.iec.ch/members\\_experts/refdocs](http://www.iec.ch/members_experts/refdocs)).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see [www.iso.org/patents](http://www.iso.org/patents)) or the IEC list of patent declarations received (see [patents.iec.ch](http://patents.iec.ch)).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see [www.iso.org/iso/foreword.html](http://www.iso.org/iso/foreword.html). In the IEC, see [www.iec.ch/understanding-standards](http://www.iec.ch/understanding-standards).

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 23090-2:2019), which has been technically revised.

The main changes compared to the previous edition are as follows:

- Multiple viewpoints have been added. Viewpoints can be used for example to provide several user-switchable camera positions to view the content or to express a storyline where the user is given the choice to select which storyline path is followed.
- Sphere-relative and viewport-relative video and image overlays have been added.
- Mesh omnidirectional video where the video is projected on an indicated set of mesh elements has been added.
- Two tiling OMAF video profiles for viewport-dependent streaming have been added.

A list of all parts in the ISO/IEC 23090 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at [www.iso.org/members.html](http://www.iso.org/members.html) and [www.iec.ch/national-committees](http://www.iec.ch/national-committees).

## Introduction

When omnidirectional media content is consumed with a head-mounted display and headphones, only the parts of the media that correspond to the user's viewing orientation are rendered, as if the user were in the spot where and when the media was captured. One of the forms of omnidirectional media applications is omnidirectional video, also known as 360° video. Omnidirectional video is typically captured by multiple cameras that cover the entire sphere or at least a large part of the sphere. Compared to traditional media application formats, the end-to-end technology for omnidirectional video (from capture to playback) is more easily fragmented due to various capturing and video projection technologies. From the capture side, there exist many different types of cameras capable of capturing 360° video, and on the playback side there are many different devices that are able to playback 360° video with different processing capabilities. To avoid fragmentation of omnidirectional media content and devices, a standardized format for omnidirectional media applications is specified in this document.

This document defines a media format that enables omnidirectional media applications, focusing on 360° video, images, and audio, as well as associated timed text. What is specified in this document includes (but is not limited to):

- a coordinate system that consists of a unit sphere and three coordinate axes, namely the X (back-to-front) axis, the Y (lateral, side-to-side) axis, and the Z (vertical, up) axis;
- projection and rectangular region-wise packing methods that may be used for conversion of a spherical video sequence or image into a two-dimensional rectangular video sequence or image, respectively;
- storage of omnidirectional media and the associated metadata using the ISO Base Media File Format (ISOBMFF) as specified in ISO/IEC 14496-12;
- storage of video or image overlays and the associated metadata using ISOBMFF;
- encapsulation, signalling, and streaming of omnidirectional media and overlays in a media streaming system, e.g. dynamic adaptive streaming over HTTP (DASH) as specified in ISO/IEC 23009-1 or MPEG media transport (MMT) as specified in ISO/IEC 23008-1;
- media profiles and presentation profiles that provide conformance points for media codecs as well as media coding and encapsulation configurations that may be used for compression, streaming, and playback of the omnidirectional media content;
- toolset brands that provide conformance points for functionalities beyond plain 360° video, images and audio.

This document is organized as follows:

- a) Clause 1 specifies the scope of this document.
- b) Clause 2 contains the normative references.
- c) Clause 3 specifies the terms, definitions, abbreviated terms, arithmetic operations, mathematical functions and other conventions used in this document.
- d) Clause 4 contains an overview of this document.
- e) Clause 5 specifies a coordinate system used in this document and the equations for the equirectangular and cubemap omnidirectional projection formats, the conversion from the local coordinate axes to the global coordinate axes, and the rectangular region-wise packing.
- f) Clause 6 specifies syntax structures that are common for fisheye video and fisheye images.
- g) Clause 7 specifies extensions to the ISOBMFF for omnidirectional media as well as for timed metadata for sphere regions. It also specifies generic extensions to ISO/IEC 14496-12 and ISO/IEC 14496-15, which may be used also for other purposes than for omnidirectional media.



- h) Clause 8 specifies extensions to DASH for omnidirectional media.
- i) Clause 9 specifies extensions to MMT for omnidirectional media.
- j) Clause 10 specifies OMAF media profiles.
- k) Clause 11 specifies OMAF presentation profiles based on some of the OMAF media profiles specified in Clause 10.
- l) Clause 12 specifies OMAF toolset brands.
- m) Annex A contains the OMAF DASH XML schema.
- n) Annex B specifies the DASH integration of all the OMAF media profiles for timed media specified in Clause 10.
- o) Annex C specifies the CMAF integration of some of the OMAF media profiles specified in Clause 10.
- p) Annex D describes some schemes for viewport-dependent omnidirectional video processing.
- q) Annex E contains some DASH MPD examples.
- r) Annex F contains some MMT signalling examples.
- s) Annex G specifies the expected OMAF player behaviour for rendering overlays.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ISO and IEC. Information may be obtained from the patent database available at [www.iso.org/patents](http://www.iso.org/patents).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those in the patent database. ISO and IEC shall not be held responsible for identifying any or all such patent rights.



# Information technology — Coded representation of immersive media —

## Part 2: Omnidirectional media format

### 1 Scope

This document specifies the omnidirectional media format for coding, storage, delivery and rendering of omnidirectional media, including video, images, audio and timed text. Omnidirectional image or video can contain graphics elements generated by computer graphics but encoded as image or video. Multiple viewpoints, each corresponding to an omnidirectional camera, are supported. The document also specifies storage and delivery of overlay images or video intended to be rendered over the omnidirectional background image or video.

### 2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 10918-1, *Information technology — Digital compression and coding of continuous-tone still images — Part 1: Requirements and guidelines*

ISO/IEC 14496-1, *Information technology — Coding of audio-visual objects — Part 1: Systems*

ISO/IEC 14496-3:2019, *Information technology — Coding of audio-visual objects — Part 3: Advanced audio coding*

Rec. ITU-T H.264 (06/19) | ISO/IEC 14496-10:2014, *Information technology — Coding of audio-visual objects — Part 10: Advanced video coding*

ISO/IEC 14496-12:2020, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 14496-14, *Information technology — Coding of audio-visual objects — Part 14, MP4 file format*

ISO/IEC 14496-15:2019, *Information technology — Coding of audio-visual objects — Part 15, Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format*

ISO/IEC 14496-30:2018, *Information technology — Coding of audio-visual objects — Part 30: Timed text and other visual overlays in ISO base media file format*

ISO/IEC 23000-19:2020, *Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media*

ISO/IEC 23000-22:2019, *Information technology — Multimedia application format — Part 22 Multi-image application format (MIAF)*

## ISO/IEC 23090-2:2021(E)

ISO/IEC 23003-3:2020, *Information technology — MPEG audio technologies — Part 3: Unified speech and audio coding*

ISO/IEC 23003-4:2020, *Information technology — MPEG audio technologies — Part 4: Dynamic range control*

ISO/IEC 23008-1:2017, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 1: MPEG media transport (MMT)*

Rec. ITU-T H.265 | ISO/IEC 23008-2:2020, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

ISO/IEC 23008-3:2019, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D audio*

ISO/IEC 23008-12, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image file format*

ISO/IEC 23009-1:2019, *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*

ISO/IEC 23091-2, *Information technology — Coding-independent code points — Part 2: Video*

ISO/IEC 23091-3, *Information technology — Coding-independent code points — Part 3: Audio*

IETF BCP 47, *Tags for Identifying Languages*

IETF Internet Standard 66, *Uniform Resource Identifier (URI): Generic Syntax*

IETF RFC 6381, *MIME Codecs and Profiles*

W3C Candidate Recommendation, *WebVTT, The Web Video Text Tracks Format*

W3C Recommendation, *TTML Profiles for Internet Media Subtitles and Captions 1.0.1 (IMSC1)*

W3C Recommendation, *XML schema part 1: Structures*

W3C Recommendation, *XML schema part 2: Datatypes*

W3C Recommendation, *XML Path Language (XPath) 2.0 (Second Edition)*