INTERNATIONAL STANDARD

ISO/IEC 23090-2

Third edition
2023-06

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

Reference number
ISO/IEC 23090-2:2023(E)

© ISO/IEC 2023



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2023

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
Website: www.iso.org

Published in Switzerland

© ISO/IEC 2023 – All rights reserved

Contents

Page

Foreword ..... **vii**

Introduction ..... **viii**

1 Scope **1**

2 Normative references **1**

3 Terms, definitions, abbreviated terms and symbols **3**
3.1 Terms and definitions 3
3.2 Abbreviated terms 11
3.3 Symbols 12
3.3.1 Arithmetic operators and mathematical functions 12
3.3.2 Order of operation precedence 13
3.3.3 Range notation 14
3.3.4 Variables 14
3.3.5 Processes 14
3.3.6 Syntax structures 15
3.3.7 Conventions for indicating the number of boxes in tables 15

4 Overview **15**
4.1 Overall architecture 15
4.2 Projected omnidirectional video/images 18
4.2.1 General 18
4.2.2 Stitching, rotation, projection, and region-wise packing 18
4.3 Fisheye omnidirectional video/images 19
4.4 Mesh omnidirectional video 20
4.5 Streaming methods for omnidirectional video 20
4.5.1 Overview 20
4.5.2 Tile-based streaming with viewport-specific author-driven binding 22
4.5.3 Tile-based streaming with free-viewport author-driven binding 22
4.5.4 Tile-based streaming with late binding 23
4.6 Additional functionalities 25
4.7 Conformance and interoperability 25
4.7.1 General 25
4.7.2 Media profiles 26
4.7.3 Presentation profiles 28
4.7.4 Toolset brands 28
4.7.5 Summary of referenceable code points 28

5 Omnidirectional video projection and region-wise packing **35**
5.1 Coordinate system 35
5.2 Omnidirectional projection formats 36
5.2.1 General 36
5.2.2 Equirectangular projection for one sample location 36
5.2.3 Cubemap projection for one sample location 37
5.3 Conversion from the local coordinate axes to the global coordinate axes 39
5.4 Region-wise packing formats 40
5.4.1 General 40
5.4.2 Conversion of one sample location for rectangular region-wise packing 41

6 Fisheye omnidirectional video **42**
6.1 General 42
6.2 The `FisheyeVideoEssentialInfoStruct()` syntax structure 42
6.2.1 Syntax 42
6.2.2 Semantics 43
6.3 The `FisheyeVideoSupplementalInfoStruct()` syntax structure 46
6.3.1 Syntax 46
6.3.2 Semantics 47

© ISO/IEC 2023 – All rights reserved

7 Omnidirectional media storage and metadata signalling in the ISOBMFF 52
7.1 Generic extensions to the ISOBMFF 52
7.1.1 Indication of a track not intended to be presented alone 52
7.1.2 Clarifications on the stereo video box 52
7.1.3 Generic sub-picture track grouping extensions 53
7.1.4 Media offset box 57
7.2 Generic extensions to ISO/IEC 14496-15 58
7.2.1 Containing of `SpatialRelationship2DDescriptionBox` for HEVC tile base track and HEVC tile tracks 58
7.3 OMAF-specific extensions to the ISOBMFF 58
7.3.1 Sync samples in timed metadata tracks 58
7.4 OMAF-specific extensions to ISO/IEC 14496-15 59
7.4.1 Coverage information box in an HEVC tile base track 59
7.5 Structures and semantics that are common for video tracks and image items 59
7.5.1 Semantics of sample locations within a decoded picture 59
7.5.2 Projection format structure 62
7.5.3 Region-wise packing structure 63
7.5.4 Rotation structure 70
7.5.5 Content coverage structure 71
7.5.6 Sphere region structure 72
7.6 Restricted video schemes for omnidirectional video 76
7.6.1 Scheme types 76
7.6.2 Projected omnidirectional video box 81
7.6.3 Fisheye omnidirectional video box 82
7.6.4 Region-wise packing box 83
7.6.5 Rotation box 83
7.6.6 Coverage information box 84
7.6.7 Mesh omnidirectional video box 84
7.6.8 Mesh box 85
7.7 Timed metadata for sphere regions 87
7.7.1 General 87
7.7.2 Sample entry 88
7.7.3 Sample format 89
7.7.4 Initial viewing orientation 89
7.7.5 Recommended viewport 91
7.7.6 Timed text sphere location metadata 94
7.8 Signalling of region-wise quality ranking 95
7.8.1 General 95
7.8.2 Spherical region-wise quality ranking 95
7.8.3 2D region-wise quality ranking 97
7.9 Storage of omnidirectional images 99
7.9.1 General 99
7.9.2 Frame packing item property 99
7.9.3 Projection format item property 100
7.9.4 Essential fisheye image item property 101
7.9.5 Supplemental fisheye image item property 102
7.9.6 Region-wise packing item property 102
7.9.7 Rotation item property 103
7.9.8 Coverage information item property 103
7.9.9 Initial viewing orientation item property 104
7.10 Storage of timed text for omnidirectional video 105
7.10.1 General 105
7.10.2 OMAF timed text configuration box 105
7.10.3 IMSC1 tracks 108
7.10.4 WebVTT tracks 108
7.11 ERP region timed metadata 109
7.11.1 General 109
7.11.2 Sample entry format 109
7.11.3 Semantics 109
7.11.4 Sample format 110

 © ISO/IEC 2023 – All rights reserved

7.11.5 Generating ERP region metadata ... 111
7.12 Storage and signalling of viewpoints for omnidirectional video and images ... 111
7.12.1 Viewpoint information structures ... 111
7.12.2 Viewpoint entity grouping ... 120
7.12.3 Timed metadata for viewpoints ... 122
7.13 Storage of omnidirectional video in sub-picture tracks ... 127
7.13.1 General ... 127
7.13.2 Projected omnidirectional video ... 127
7.13.3 Indication of composition pictures being packed pictures or projected pictures ... 128
7.13.4 Fisheye omnidirectional video ... 128
7.14 Storage and signalling of overlays for omnidirectional video and images ... 129
7.14.1 General ... 129
7.14.2 Overlay structure ... 131
7.14.3 Overlay control structures ... 132
7.14.4 Overlay configuration box ... 143
7.14.5 Overlay item property ... 143
7.14.6 Overlay timed metadata track ... 144
7.14.7 Entity groups ... 145
7.14.8 Overlay alpha auxiliary image ... 148
7.15 Signalling of viewing space information ... 149
7.15.1 General ... 149
7.15.2 Viewing space structure ... 149
7.15.3 Viewing space box ... 152
7.15.4 Viewing space item property ... 152
7.15.5 Time varying immersive viewing space signalling ... 152
7.16 Mapping of rectangular regions to the 3D mesh ... 153
7.16.1 General ... 153
7.16.2 Tile mesh sample grouping ... 153
7.16.3 Rectangular region structure ... 155
7.16.4 Projection of a sample location onto the 3D mesh ... 156

8 Omnidirectional media encapsulation and signalling in DASH ... 157
8.1 Architecture of DASH delivery in OMAF ... 157
8.2 Usage of DASH in OMAF ... 159
8.2.1 General ... 159
8.2.2 Signalling of stereoscopic frame packing ... 159
8.2.3 Carriage of timed metadata ... 159
8.2.4 Associating Adaptation Sets or Representations with each other ... 160
8.3 DASH MPD descriptors for omnidirectional media in the namespace "urn:mpeg:mpegI:omaf:2017" ... 161
8.3.1 XML namespace and schema ... 161
8.3.2 Signalling of projection type information ... 161
8.3.3 Signalling of region-wise packing type ... 162
8.3.4 Signalling of content coverage ... 163
8.3.5 Signalling of spherical region-wise quality ranking ... 166
8.3.6 Signalling of 2D region-wise quality ranking ... 172
8.3.7 Signalling of fisheye omnidirectional video ... 177
8.4 Carriage of images ... 177
8.4.1 General ... 177
8.4.2 Format and constraints for Segments ... 178
8.5 DASH MPD descriptors for omnidirectional media in the namespace "urn:mpeg:mpegI:omaf:2020" ... 178
8.5.1 XML namespace and schema ... 178
8.5.2 Signalling of association ... 178
8.5.3 Signalling of viewpoints ... 180
8.5.4 Signalling of sub-picture composition identifier and its attributes ... 187
8.5.5 Signalling of overlays ... 188
8.5.6 Entity to group descriptor ... 190
8.5.7 Content component attribute for Representation ... 192
8.6 Segment formats ... 192

© ISO/IEC 2023 – All rights reserved

8.6.1 Initialization Segment for OMAF base track ... 192
8.6.2 Tile Index Segment ... 196
8.6.3 Tile Data Segment ... 197

9 Omnidirectional media encapsulation and signalling in MMT ... 198
9.1 Architecture of MMT delivery in OMAF ... 198
9.2 OMAF signalling in MPEG composition information ... 199
9.3 VR application-specific MMT signalling ... 199
9.3.1 General ... 199
9.3.2 MMT signalling ... 200

10 Media profiles ... 216
10.1 Video profiles ... 216
10.2 Audio profiles ... 250
10.3 Image profiles ... 259
10.4 Timed text profiles ... 266

11 Presentation profiles ... 267
11.1 OMAF viewport-independent baseline presentation profile ... 267
11.1.1 General ... 267
11.1.2 ISO Base Media File Format constraints ... 267
11.2 OMAF viewport-dependent baseline presentation profile ... 268
11.2.1 General ... 268
11.2.2 ISO Base Media File Format constraints ... 268

12 OMAF toolset brands ... 268
12.1 Overlay toolset brand ... 268
12.1.1 Overview ... 268
12.1.2 ISO Base Media File Format constraints ... 268
12.1.3 OMAF player operation ... 269
12.2 Viewpoint toolset brand ... 269
12.2.1 Overview ... 269
12.2.2 ISO Base Media File Format constraints ... 269
12.2.3 OMAF player operation ... 269
12.3 Non-linear storyline toolset brand ... 269
12.3.1 Overview ... 269
12.3.2 ISO Base Media File Format constraints ... 269
12.3.3 OMAF player operation ... 270

Annex A (normative) OMAF DASH schema ... 271

Annex B (normative) DASH integration of media profiles ... 275

Annex C (normative) CMAF integration of media profiles ... 288

Annex D (informative) Viewport-dependent omnidirectional video processing ... 292

Annex E (informative) DASH MPD examples ... 335

Annex F (informative) MMT signalling examples ... 340

Annex G (normative) Expected behaviour of OMAF player ... 342

Bibliography ... 352

 © ISO/IEC 2023 – All rights reserved

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 or 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) or the IEC list of patent declarations received (see 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. In the IEC, see 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 third edition cancels and replaces the second edition (ISO/IEC 23090-2:2021), which has been technically revised.

The main changes are as follows:

— The following profiles for the Versatile Video Coding (VVC) standard (Rec. ITU-T H.266 | ISO/IEC 23090-3) have been added:

  — VVC-based simple tiling OMAF video profile,

  — VVC-based viewport-independent OMAF video profile,

  — OMAF VVC image profile, and

  — CMAF media profile for the VVC-based viewport-independent OMAF video profile.

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 and www.iec.ch/national-committees.

© ISO/IEC 2023 – All rights reserved

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.

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.

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.

 © ISO/IEC 2023 – All rights reserved

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 14496-1:2010, *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 (08/21) | ISO/IEC 14496-10:2020, *Information technology — Coding of audio-visual objects — Part 10: Advanced video coding*

ISO/IEC 14496-12:2022, *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 14496-15:2022, *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 2023 – All rights reserved

ISO/IEC 23000-19:—[1], *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)*

Rec. ITU-T H.274 | ISO/IEC 23002-7, *Information technology — MPEG video technologies — Part 7: Versatile supplemental enhancement information messages for coded video bitstreams*

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 (11/19) | ISO/IEC 23008-2:2020, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

ISO/IEC 23008-3:2022, *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 23009-1:2022, *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*

Rec. ITU-T H.266 (08/20) | ISO/IEC 23090-3:2021, *Information technology — Coded representation of immersive media — Part 3: Versatile video coding*

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*

[1] Under preparation. Stage at the time of publication: ISO/IEC DIS 23000-19:2022.

© ISO/IEC 2023 – All rights reserved