

ISO/IEC 23090-2:2021-07 (E)

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

Contents	Page
Foreword.....	vii
Introduction.....	viii
1 Scope	1
2 Normative references	1
3 Terms, definitions, abbreviated terms and conventions	2
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
4 Overview.....	13
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
5 Omnidirectional video projection and region-wise packing	32
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
6 Fisheye omnidirectional video.....	39
6.1 General.....	39
6.2 The <code>FisheyeVideoEssentialActionInfoStruct()</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

7	Omni-directional media storage and metadata signalling in the ISO/BMFF	48
7.1	Generic extensions to the ISO/BMFF.....	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 <code>SpatialRelationship2DDescriptionBox</code> for HEVC tile base track and HEVC tile tracks.....	56
7.3	OMAF-specific extensions to the ISO/BMFF	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 omni-directional video	73
7.6.1	Scheme types.....	73
7.6.2	Projected omni-directional video box.....	78
7.6.3	Fisheye omni-directional 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 omni-directional 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 omni-directional 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 omni-directional 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
8	Omnidirectional media encapsulation and signalling in DASH	153
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
9	Omnidirectional media encapsulation and signalling in MMT	194
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
10	Media profiles.....	212
10.1	Video profiles.....	212
10.2	Audio profiles	243
10.3	Image profiles.....	252
10.4	Timed text profiles	257
11	Presentation profiles	258
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
12	OMAF toolset brands.....	260
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
	Annex A (normative) OMAF DASH schema.....	262
	Annex B (normative) DASH integration of media profiles	266
	Annex C (normative) CMAF integration of media profiles	279
	Annex D (informative) Viewport-dependent omnidirectional video processing.....	282
	Annex E (informative) DASH MPD examples.....	316
	Annex F (informative) MMT signalling examples.....	320
	Annex G (normative) Expected behaviour of OMAF player	322
	Bibliography.....	332