ISO/IEC 23090-2:2023-06 (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 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
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

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

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