

ISO/IEC 14496-12:2026-04 (E)

Information technology - Coding of audio-visual objects - Part 12: ISO base media file format

Contents

Page

Foreword.....	x
Introduction.....	xi
1 Scope.....	1
2 Normative references.....	1
3 Terms, definitions and abbreviated terms.....	2
3.1 Terms and definitions.....	2
3.2 Abbreviated terms.....	8
4 ISO base media file organization.....	9
4.1 File structure.....	9
4.2 Binary structure.....	9
4.2.1 Syntax conventions.....	9
4.2.2 Box definitions.....	10
4.2.3 Extensibility of box definitions.....	11
4.2.4 Data types and fields.....	12
5 Typed ISO base media file organization.....	13
5.1 File structure.....	13
5.2 File-type box.....	14
5.2.1 Definition.....	14
5.2.2 Syntax.....	14
5.2.3 Semantics.....	14
5.3 Extended type box.....	15
5.3.1 Definition.....	15
5.3.2 Syntax.....	15
5.3.3 Semantics.....	15
6 Common specifications for movie files, item files and segment files.....	16
6.1 Data offsets.....	16
6.2 Box order.....	17
7 Movie file organization.....	21
7.1 Files, segments, and streams.....	21
7.2 Presentation structure.....	22
7.2.1 Box structure of a presentation.....	22
7.2.2 Meta data and media data.....	22
7.3 Structure-data.....	22
7.3.1 Box.....	22
7.3.2 Interpretation of matrix values.....	22
7.3.3 URIs as type indicators.....	23
7.3.4 Box order.....	23
7.4 Time structure overview.....	24
7.5 Identifiers.....	24
7.6 Uniform resource locators (URLs).....	24
8 Box structures.....	25
8.1 File structure and general boxes.....	25
8.1.1 Media data box.....	25
8.1.2 Free space box.....	25
8.1.3 Progressive download information box.....	26
8.1.4 Identified media data box.....	26

8.2	Movie structure	27
8.2.1	Movie box	27
8.2.2	Movie header box	27
8.3	Track structure	29
8.3.1	Track box	29
8.3.2	Track header box	29
8.3.3	Track reference box	32
8.3.4	Track group box	34
8.3.5	Track type box	36
8.4	Track media structure	37
8.4.1	Media box	37
8.4.2	Media header box	37
8.4.3	Handler box	38
8.4.4	Media information box	39
8.4.5	Media information header boxes	39
8.4.6	Extended language tag	40
8.5	Sample tables	40
8.5.1	Sample table box	40
8.5.2	Sample description box	41
8.5.3	Degradation priority box	43
8.5.4	Sample scale box	43
8.6	Track time structures	43
8.6.1	Time to sample boxes	43
8.6.2	Sync sample box	49
8.6.3	Shadow sync	50
8.6.4	Independent and disposable samples box	52
8.6.5	Edit box	53
8.6.6	Edit list box	54
8.7	Track data layout structures	56
8.7.1	Data information box	56
8.7.2	Data reference box	57
8.7.3	Sample size boxes	59
8.7.4	Sample to chunk box	60
8.7.5	Chunk offset box	61
8.7.6	Padding bits box	62
8.7.7	Sub-sample information box	62
8.7.8	Sample auxiliary information sizes box	65
8.7.9	Sample auxiliary information offsets box	67
8.7.10	Sub-sample Reference Table Box	68
8.8	Movie fragments	69
8.8.1	Movie extends box	69
8.8.2	Movie extends header box	70
8.8.3	Track extends box	71
8.8.4	Movie fragment box	72
8.8.5	Movie fragment header box	72
8.8.6	Track fragment box	73
8.8.7	Track fragment header box	73
8.8.8	Track fragment run box	75
8.8.9	Movie fragment random access box	76
8.8.10	Track fragment random access box	77
8.8.11	Movie fragment random access offset box	78
8.8.12	Track fragment decode time box	79
8.8.13	Level assignment box	80
8.8.14	Sample auxiliary information in movie fragments	81
8.8.15	Track extension properties box	82
8.8.16	Alternative startup sequence properties box	82
8.8.17	Metadata and user data in movie fragments	83
8.8.18	Redundant Sample Original Timing	83
8.9	Sample group structures	85
8.9.1	Overview	85
8.9.2	Sample to group box	85
8.9.3	Sample group description box	87

8.9.4	Representation of group structures in movie fragments	89
8.9.5	Compact sample to group box	90
8.10	User data	92
8.10.1	User data box	92
8.10.2	Copyright box	93
8.10.3	Track selection box	93
8.10.4	Track kind	95
8.10.5	Label box	95
8.11	Metadata support	96
8.11.1	MetaBox	96
8.11.2	XML boxes	98
8.11.3	Item location box	98
8.11.4	Primary item box	102
8.11.5	Item protection box	102
8.11.6	Item information box	103
8.11.7	Additional metadata container box	105
8.11.8	Metabox Relation box	105
8.11.9	URL forms for MetaBoxes	105
8.11.10	Static metadata	106
8.11.11	Item data box	107
8.11.12	Item reference box	107
8.11.13	Auxiliary video metadata	108
8.11.14	Item properties box	108
8.11.15	Brand item property	110
8.11.16	Handler property	110
8.11.17	T.35 item	111
8.12	File delivery format support	111
8.12.1	Overview	111
8.12.2	FD item information box	112
8.12.3	File partition box	113
8.12.4	FEC reservoir box	114
8.12.5	FD session group box	115
8.12.6	Group ID to name box	116
8.12.7	File reservoir box	116
8.13	Sub tracks	117
8.13.1	Overview	117
8.13.2	Backward compatibility	117
8.13.3	Sub track box	118
8.13.4	Sub track information box	118
8.13.5	Sub track definition box	119
8.13.6	Sub track sample group box	120
8.14	Segments	120
8.14.1	Overview	120
8.14.2	Segment type box	120
8.14.3	Segment index box	121
8.14.4	Subsegment index box	124
8.14.5	Producer reference time box	126
8.15	Entity grouping	127
8.15.1	General	127
8.15.2	Groups list box	128
8.15.3	Entity to group box	128
8.15.4	Entity group definitions	129
8.16	Compressed boxes	131
8.16.1	Overview and processing	131
8.16.2	Processing model	131
8.16.3	General syntax	133
8.16.4	General semantics	133
8.16.5	Original file-type box	133
8.16.6	Compressed movie box	134
8.16.7	Compressed movie fragment box	134
8.16.8	Compressed segment index box	134
8.16.9	Compressed subsegment index box	135

9	Hint track formats	135
9.1	RTP and SRTP hint track format.....	135
9.1.1	Overview.....	135
9.1.2	Sample entry format.....	136
9.1.3	Sample format.....	138
9.1.4	SDP information.....	140
9.1.5	Statistical information.....	141
9.2	ALC/LCT and FLUTE hint track format.....	141
9.2.1	Overview.....	141
9.2.2	Design principles.....	142
9.2.3	Sample entry format.....	143
9.2.4	Sample format.....	144
9.3	MPEG-2 transport hint track format.....	147
9.3.1	Overview.....	147
9.3.2	Design principles.....	147
9.3.3	Sample entry format.....	149
9.3.4	Sample format.....	151
9.3.5	Protected MPEG 2 transport stream hint track.....	153
9.4	RTP, RTCP, SRTP and SRTCP reception hint tracks.....	153
9.4.1	RTP reception hint track.....	153
9.4.2	RTCP reception hint track.....	156
9.4.3	SRTP reception hint track.....	158
9.4.4	SRTCP reception hint tracks.....	159
9.4.5	Protected RTP reception hint track.....	160
9.4.6	Recording procedure.....	160
9.4.7	Parsing procedure.....	161
10	Sample groups	161
10.1	Random access recovery points.....	161
10.1.1	Definition.....	161
10.1.2	Syntax.....	161
10.1.3	Semantics.....	161
10.2	Rate share groups.....	162
10.2.1	Overview.....	162
10.2.2	Rate share sample group description entry.....	163
10.2.3	Relationship between tracks.....	164
10.2.4	Bitrate allocation.....	164
10.3	Alternative startup sequences.....	165
10.3.1	Definition.....	165
10.3.2	Syntax.....	165
10.3.3	Semantics.....	166
10.3.4	Examples.....	166
10.4	Random access point (RAP) sample group.....	168
10.4.1	Definition.....	168
10.4.2	Syntax.....	169
10.4.3	Semantics.....	169
10.5	Temporal level sample group.....	169
10.5.1	Definition.....	169
10.5.2	Syntax.....	169
10.5.3	Semantics.....	169
10.6	Stream access point sample group.....	169
10.6.1	Definition.....	169
10.6.2	Syntax.....	170
10.6.3	Semantics.....	170
10.7	Sample-to-item sample group.....	170
10.7.1	Definition.....	170
10.7.2	Syntax.....	170
10.7.3	Semantics.....	170
10.8	Dependent random access point (DRAP) sample group.....	171

10.8.1	Definition	171
10.8.2	Syntax	171
10.8.3	Semantics	171
10.9	Pixel Aspect Ratio Sample Grouping	172
10.9.1	Definition	172
10.9.2	Syntax	172
10.9.3	Semantics	172
10.10	Clean Aperture Sample Grouping	172
10.10.1	Definition	172
10.10.2	Syntax	172
10.10.3	Semantics	172
10.11	EDRAP sample group	173
10.11.1	Definition	173
10.11.2	Syntax	173
10.11.3	Semantics	173
10.12	Essential descriptions hierarchy sample grouping	174
10.12.1	Definition	174
10.12.2	Syntax	175
10.12.3	Semantics	175
10.13	T.35 sample group	175
10.13.1	Definition	175
10.13.2	Syntax	176
10.13.3	Semantics	176
11	Derived file formats	176
12	Media-specific definitions	177
12.1	Video media	177
12.1.1	Media handler	177
12.1.2	Video media header	177
12.1.3	Sample entry	178
12.1.4	Pixel aspect ratio and clean aperture	178
12.1.5	Colour information	180
12.1.6	Content light level	181
12.1.7	Mastering display colour volume	181
12.1.8	Content colour volume	181
12.1.9	Ambient viewing environment	182
12.2	Audio media	182
12.2.1	Media handler	182
12.2.2	Sound media header	183
12.2.3	Sample entry	183
12.2.4	Channel layout	185
12.2.5	Downmix instructions	187
12.2.6	DRC information	191
12.2.7	Audio stream loudness	191
12.2.8	Audio rendering indication box	194
12.2.9	Audio element box	194
12.2.10	Audio element description box	195
12.2.11	Audio element positioning interactivity polar box	195
12.2.12	Audio element prominence interactivity box	197
12.2.13	Audio element selection box	197
12.2.14	Audio element selection description box	198
12.3	Metadata media	198
12.3.1	Media handler	198
12.3.2	Media header	198
12.3.3	Sample entry	198
12.4	Hint media	200
12.4.1	Overview	200
12.4.2	Media handler	201
12.4.3	Hint media header	201

12.4.4	Sample entry	202
12.5	Text media	202
12.5.1	Media handler	202
12.5.2	Media header	202
12.5.3	Sample entry	202
12.6	Subtitle media	203
12.6.1	Media handler	203
12.6.2	Subtitle media header	203
12.6.3	Sample entry	203
12.7	Font media	204
12.7.1	Media handler	204
12.7.2	Media header	204
12.7.3	Sample entry	204
12.8	Multiplexed timed metadata tracks	204
12.8.1	General	204
12.8.2	Overall design	205
12.8.3	Sample format	205
12.8.4	Sample entry format	205
12.8.5	Defined formats	208
12.9	Volumetric visual media	209
12.9.1	Media handler	209
12.9.2	Media header	209
12.9.3	Sample entry	209
12.9.4	Sample format	210
12.10	Haptic media	210
12.10.1	Media handler	210
12.10.2	Media header	210
12.10.3	Sample entry	210
12.10.4	Sample format	210
13	Transformed media tracks	210
13.1	General	210
13.2	Multiple transformations for a single transformed media track	211
13.3	Determining the untransformed sample entry type	211
13.3.1	General	211
13.3.2	Example for protected and restricted media (informative)	211
13.3.3	The 'codecs' MIME parameter for a transformed media track	212
13.4	Support for protected streams	212
13.4.1	Overview	212
13.4.2	Protection scheme information box	213
13.4.3	Original format box	214
13.4.4	IPMPInfoBox	214
13.4.5	IPMP control box	214
13.4.6	Scheme type box	214
13.4.7	Scheme information box	215
13.4.8	Scramble Scheme Information Box	215
13.5	Restricted media tracks	216
13.5.1	General	216
13.5.2	Restricted sample entry transformation	216
13.5.3	Restricted scheme information box	217
13.5.4	Scheme for stereoscopic video arrangements	217
13.5.5	Compatible scheme type box	220
13.5.6	Sample-packed tracks	220
13.6	Support for incomplete tracks	221
13.6.1	General	221
13.6.2	Transformation	222
13.6.3	Complete track information box	222
	Annex A (informative) Overview and tutorial	223

Annex B (informative) **Guidance on deriving from this document**.....231
Annex C (normative) **Fragment identifiers for ISO base media resources** 240
Annex D (informative) **Management of extension code-points**.....241
Annex E (normative) **File format brands**.....242
Annex F (normative) **MIME type registration of segments**255
Annex G (informative) **URI-labelled metadata forms**.....256
Annex H (informative) **Processing of RTP streams and reception hint tracks**258
Annex I (normative) **Stream access points**.....273
Annex J (informative) **Segment index examples**276
Annex K (normative) **Use of IETF RFC 6381 for ISOBMFF files**281
Bibliography.....284