
**Information technology — Coding of audio-
visual objects —**

**Part 12:
ISO base media file format**

*Technologies de l'information — Codage des objets audiovisuels —
Partie 12: Format ISO de base pour les fichiers médias*

Reference number

ISO/IEC 14496-12:2015(E)



Copyright notice

This ISO document is a Draft International Standard and is copyright protected by ISO. Except as permitted under the applicable laws of the user's country, neither this ISO draft nor any extract from it may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, photocopying, recording or otherwise, without prior written permission being secured.

Requests for permission to reproduce should be addressed to either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH 1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Reproduction may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

Contents

Page

1	Scope	1
2	Normative references	1
3	Terms, definitions, and abbreviated terms	3
3.1	Terms and definitions	3
3.2	Abbreviated terms	5
4	Object structured File Organization	5
4.1	File Structure	5
4.2	Object Structure	6
4.3	File Type Box	7
5	Design Considerations	8
5.1	Usage	8
5.1.1	Introduction	8
5.1.2	Interchange	8
5.1.3	Content Creation	8
5.1.4	Preparation for streaming	9
5.1.5	Local presentation	9
5.1.6	Streamed presentation	10
5.2	Design principles	10
6	ISO Base Media File organization	11
6.1	Presentation structure	11
6.1.1	File Structure	11
6.1.2	Object Structure	11
6.1.3	Meta Data and Media Data	12
6.1.4	Track Identifiers	12
6.2	Metadata Structure (Objects)	12
6.2.1	Box	12
6.2.2	Data Types and fields	12
6.2.3	Box Order	13
6.2.4	URLs as type indicators	16
6.3	Brand Identification	17
7	Streaming Support	17
7.1	Handling of Streaming Protocols	17
7.2	Protocol 'hint' tracks	18
7.3	Hint Track Format	18
8	Box Structures	20
8.1	File Structure and general boxes	20
8.1.1	Media Data Box	20
8.1.2	Free Space Box	20
8.1.3	Progressive Download Information Box	21
8.2	Movie Structure	21
8.2.1	Movie Box	21

8.2.2 — Movie Header Box	22
8.3 — Track Structure	23
8.3.1 — Track Box	23
8.3.2 — Track Header Box	23
8.3.3 — Track Reference Box	25
8.3.4 — Track Group Box	27
8.4 — Track Media Structure	28
8.4.1 — Media Box	28
8.4.2 — Media Header Box	28
8.4.3 — Handler Reference Box	29
8.4.4 — Media Information Box	30
8.4.5 — Media Information Header Boxes	30
8.4.6 — Extended language tag	31
8.5 — Sample Tables	31
8.5.1 — Sample Table Box	31
8.5.2 — Sample Description Box	32
8.5.3 — Degradation Priority Box	34
8.5.4 — Sample Scale Box	34
8.6 — Track Time Structures	34
8.6.1 — Time to Sample Boxes	34
8.6.2 — Sync Sample Box	39
8.6.3 — Shadow Sync Sample Box	40
8.6.4 — Independent and Disposable Samples Box	41
8.6.5 — Edit Box	43
8.6.6 — Edit List Box	43
8.7 — Track Data Layout Structures	45
8.7.1 — Data Information Box	45
8.7.2 — Data Reference Box	45
8.7.3 — Sample Size Boxes	46
8.7.4 — Sample To Chunk Box	48
8.7.5 — Chunk Offset Box	48
8.7.6 — Padding Bits Box	49
8.7.7 — Sub-Sample Information Box	50
8.7.8 — Sample Auxiliary Information Sizes Box	51
8.7.9 — Sample Auxiliary Information Offsets Box	53
8.8 — Movie Fragments	54
8.8.1 — Movie Extends Box	54
8.8.2 — Movie Extends Header Box	54
8.8.3 — Track Extends Box	55
8.8.4 — Movie Fragment Box	56
8.8.5 — Movie Fragment Header Box	56
8.8.6 — Track Fragment Box	57
8.8.7 — Track Fragment Header Box	57
8.8.8 — Track Fragment Run Box	58
8.8.9 — Movie Fragment Random Access Box	60
8.8.10 — Track Fragment Random Access Box	60

8.8.11	Movie Fragment Random Access Offset Box	61
8.8.12	Track fragment decode time	62
8.8.13	Level Assignment Box	63
8.8.14	Sample Auxiliary Information in Movie Fragments	65
8.8.15	Track Extension Properties Box	65
8.8.16	Alternative Startup Sequence Properties Box	66
8.8.17	Metadata and user data in movie fragments	66
8.9	Sample Group Structures	67
8.9.1	Introduction	67
8.9.2	Sample to Group Box	68
8.9.3	Sample Group Description Box	69
8.9.4	Representation of group structures in Movie Fragments	70
8.10	User Data	71
8.10.1	User Data Box	71
8.10.2	Copyright Box	72
8.10.3	Track Selection Box	72
8.10.4	Track kind	74
8.11	Metadata Support	75
8.11.1	The Meta box	75
8.11.2	XML Boxes	76
8.11.3	The Item Location Box	77
8.11.4	Primary Item Box	80
8.11.5	Item Protection Box	80
8.11.6	Item Information Box	81
8.11.7	Additional Metadata Container Box	83
8.11.8	Metabox Relation Box	84
8.11.9	URL Forms for meta boxes	85
8.11.10	Static Metadata	85
8.11.11	Item Data Box	86
8.11.12	Item Reference Box	87
8.11.13	Auxiliary video metadata	88
8.12	Support for Protected Streams	88
8.12.1	Protection Scheme Information Box	89
8.12.2	Original Format Box	90
8.12.3	IPMPInfoBox	90
8.12.4	IPMP Control Box	90
8.12.5	Scheme Type Box	90
8.12.6	Scheme Information Box	91
8.13	File Delivery Format Support	91
8.13.1	Introduction	91
8.13.2	FD Item Information Box	92
8.13.3	File Partition Box	92
8.13.4	FEC Reservoir Box	94
8.13.5	FD Session Group Box	95
8.13.6	Group ID to Name Box	96
8.13.7	File Reservoir Box	96

8.14	Sub tracks	97
8.14.1	Introduction	97
8.14.2	Backward compatibility	97
8.14.3	Sub Track box	98
8.14.4	Sub Track Information box	98
8.14.5	Sub Track Definition box	100
8.14.6	Sub Track Sample Group box	100
8.15	Post-decoder requirements on media	100
8.15.1	General	100
8.15.2	Transformation	101
8.15.3	Restricted Scheme Information box	102
8.15.4	Scheme for stereoscopic video arrangements	102
8.16	Segments	104
8.16.1	Introduction	104
8.16.2	Segment Type Box	104
8.16.3	Segment Index Box	105
8.16.4	Subsegment Index Box	109
8.16.5	Producer Reference Time Box	111
8.17	Support for Incomplete Tracks	112
8.17.1	General	112
8.17.2	Transformation	113
8.17.3	Complete Track Information Box	114
9	Hint Track Formats	114
9.1	RTP and SRTP Hint Track Format	114
9.1.1	Introduction	114
9.1.2	Sample Description Format	115
9.1.3	Sample Format	117
9.1.4	SDP Information	119
9.1.5	Statistical Information	120
9.2	ALC/LCT and FLUTE Hint Track Format	121
9.2.1	Introduction	121
9.2.2	Design principles	122
9.2.3	Sample Description Format	123
9.2.4	Sample Format	124
9.3	MPEG-2 Transport Hint Track Format	127
9.3.1	Introduction	127
9.3.2	Design Principles	128
9.3.3	Sample Description Format	130
9.3.4	Sample Format	132
9.3.5	Protected MPEG 2 Transport Stream Hint Track	134
9.4	RTP, RTCP, SRTP and SRTCP Reception Hint Tracks	134
9.4.1	RTP Reception Hint Track	134
9.4.2	RTCP Reception Hint Track	138
9.4.3	SRTP Reception Hint Track	140
9.4.4	SRTCP Reception Hint Tracks	142

9.4.5	Protected RTP Reception Hint Track	143
9.4.6	Recording Procedure	143
9.4.7	Parsing Procedure	143
10	Sample Groups	143
10.1	Random Access Recovery Points	143
10.2	Rate Share Groups	144
10.2.1	Introduction	144
10.2.2	Rate Share Sample Group Entry	146
10.2.3	Relationship between tracks	147
10.2.4	Bitrate allocation	147
10.3	Alternative Startup Sequences	148
10.3.4	Examples	149
10.4	Random Access Point (RAP) Sample Grouping	151
10.5	Temporal level sample grouping	152
10.6	Stream access point sample group	152
11	Extensibility	153
11.1	Objects	153
11.2	Storage formats	154
11.3	Derived File formats	154
12	Media-specific definitions	155
12.1	Video media	155
12.1.1	Media handler	155
12.1.2	Video media header	155
12.1.3	Sample entry	156
12.1.4	Pixel Aspect Ratio and Clean Aperture	156
12.1.5	Colour information	158
12.2	Audio media	159
12.2.1	Media handler	159
12.2.2	Sound media header	159
12.2.3	Sample entry	160
12.2.4	Channel layout	162
12.2.5	Downmix Instructions	163
12.2.6	DRC Information	165
12.2.7	Audio stream loudness	165
12.3	Metadata media	167
12.3.1	Media handler	167
12.3.2	Media header	167
12.3.3	Sample entry	167
12.4	Hint media	169
12.4.1	Media handler	169
12.4.2	Hint media header	169
12.4.3	Sample entry	170
12.5	Text media	170
12.5.1	Media handler	170
12.5.2	Media header	170

12.5.3—Sample entry	170
12.6—Subtitle media	171
12.6.1—Media handler	171
12.6.2—Subtitle media header	171
12.6.3—Sample entry	171
12.7—Font media	172
12.7.1—Media handler	172
12.7.2—Media header	172
12.7.3—Sample entry	172
12.8—Transformed media	172
Annex A (informative) Overview and Introduction	173
A.1—Section Overview	173
A.2—Core Concepts	173
A.3—Physical structure of the media	174
A.4—Temporal structure of the media	174
A.5—Interleave	175
A.6—Composition	175
A.7—Random access	175
A.8—Fragmented movie files	176
Annex B (void)	178
Annex C (informative) Guidelines on deriving from this specification	179
C.1—Introduction	179
C.2—General Principles	179
C.2.1—General	179
C.2.2—Base layer operations	180
C.3—Boxes	180
C.4—Brand Identifiers	181
C.4.1—Introduction	181
C.4.2—Usage of the Brand	181
C.4.3—Introduction of a new brand	182
C.4.4—Player Guideline	182
C.4.5—Authoring Guideline	182
C.4.6—Example	183
C.5—Storage of new media types	183
C.6—Use of Template fields	183
C.7—Tracks	184
C.7.1—Data Location	184
C.7.2—Time	184
C.7.3—Media Types	185
C.7.4—Coding Types	185
C.7.5—Sub-sample information	185
C.7.6—Sample Dependency	185
C.7.7—Sample Groups	185
C.7.8—Track-level	186
C.7.9—Protection	186

C.8	Construction of fragmented movies	186
C.9	Meta-data	187
C.10	Registration	187
C.11	Guidelines on the use of sample groups, timed metadata tracks, and sample auxiliary information	187
Annex D	(informative) Registration Authority	190
D.1	Code points to be registered	190
D.2	Procedure for the request of an MPEG-4 registered identifier value	191
D.3	Responsibilities of the Registration Authority	191
D.4	Contact information for the Registration Authority	191
D.5	Responsibilities of Parties Requesting a RID	192
D.6	Appeal Procedure for Denied Applications	192
D.7	Registration Application Form	192
D.7.1	Contact Information of organization requesting a RID	192
D.7.2	Request for a specific RID	193
D.7.3	Short description of RID that is in use and date system was implemented	193
D.7.4	Statement of an intention to apply the assigned RID	193
D.7.5	Date of intended implementation of the RID	193
D.7.6	Authorized representative	193
D.7.7	For official use of the Registration Authority	194
Annex E	(normative) File format brands	195
E.1	Introduction	195
E.2	The 'isom' brand	196
E.3	The 'avc1' brand	197
E.4	The 'iso2' brand	197
E.5	The 'mp71' brand	198
E.6	The 'iso3' brand	198
E.7	The 'iso4' brand	199
E.8	The 'iso5' brand	199
E.9	The 'iso6' brand	200
E.10	The 'iso7' brand	200
E.11	The 'iso8' brand	201
E.12	The 'iso9' brand	201
Annex F	(void)	202
Annex G	(informative) URI-labelled metadata forms	203
G.1	UUID-labelled metadata	203
G.2	ISO OID-labelled metadata	203
G.3	SMPTE-labelled metadata	204
Annex H	(informative) Processing of RTP streams and reception hint tracks	205
H.1	Introduction	205
H.1.1	Overview	205
H.1.2	Structure	205
H.1.3	Terms and definitions	205
H.2	Synchronization of RTP streams	205

H.3—Recording of RTP streams	206
H.3.1—Introduction	206
H.3.2—Compensation for unequal starting for position of received RTP streams	209
H.3.3—Recording of SDP	210
H.3.4—Creation of a sample within an RTP reception hint track	210
H.3.5—Representation of RTP timestamps	211
H.3.6—Recording operations to facilitate inter-stream synchronization in playback	214
H.3.7—Representation of reception times	216
H.3.8—Creation of media samples	216
H.3.9—Creation of hint samples referring to media samples	217
H.4—Playing of recorded RTP streams	217
H.4.1—Introduction	217
H.4.2—Preparation for the playback	218
H.4.3—Decoding of a sample within an RTP reception hint track	218
H.4.4—Lip-synchronization	219
H.4.5—Random access	220
H.5—Re-sending recorded RTP streams	221
H.5.1—Introduction	221
H.5.2—Re-sending RTP packets	222
H.5.3—RTCP Processing	223
Annex I (normative) Stream Access Points	224
I.1—Introduction	224
I.2—SAP properties	224
I.2.1—General	224
I.2.2—SAP properties for layers	225
I.3—SAP types	226
Annex J (normative) MIME Type Registration of Segments	227
J.1—Introduction	227
J.2—Registration	227
Annex K : Segment Index Examples (informative)	228
K.1—Introduction	228
K.2—Examples	228
K.2.1—Simple one-level indexing	228
K.2.2—Hierarchical	228
K.2.3—Daisy-chain	229
K.2.4—Combination hierarchical and daisy-chain	230
1 Scope	1
2 Normative references	1
3 Terms, definitions, and abbreviated terms	3
3.1 Terms and definitions	3
3.2 Abbreviated terms	5
4 Object-structured File Organization	6
4.1 File Structure	6

4.2	Object Structure.....	6
4.3	File Type Box	7
5	Design Considerations	8
5.1	Usage.....	8
5.1.1	Introduction	8
5.1.2	Interchange.....	8
5.1.3	Content Creation.....	9
5.1.4	Preparation for streaming.....	10
5.1.5	Local presentation	10
5.1.6	Streamed presentation.....	10
5.2	Design principles	10
6	ISO Base Media File organization	11
6.1	Presentation structure.....	11
6.1.1	File Structure	11
6.1.2	Object Structure.....	12
6.1.3	Meta Data and Media Data	12
6.1.4	Track Identifiers.....	12
6.2	Metadata Structure (Objects)	12
6.2.1	Box	12
6.2.2	Data Types and fields	12
6.2.3	Box Order	14
6.2.4	URLs as type indicators.....	16
6.3	Brand Identification.....	17
7	Streaming Support	17
7.1	Handling of Streaming Protocols	17
7.2	Protocol 'hint' tracks.....	18
7.3	Hint Track Format	18
8	Box Structures	20
8.1	File Structure and general boxes.....	20
8.1.1	Media Data Box.....	20
8.1.2	Free Space Box.....	20
8.1.3	Progressive Download Information Box.....	21
8.2	Movie Structure	21
8.2.1	Movie Box.....	21
8.2.2	Movie Header Box	22
8.3	Track Structure.....	23
8.3.1	Track Box	23
8.3.2	Track Header Box.....	23
8.3.3	Track Reference Box	25
8.3.4	Track Group Box.....	27
8.4	Track Media Structure	28
8.4.1	Media Box.....	28
8.4.2	Media Header Box	28
8.4.3	Handler Reference Box.....	29

8.4.4	Media Information Box	30
8.4.5	Media Information Header Boxes	30
8.4.6	Extended language tag	31
8.5	Sample Tables	31
8.5.1	Sample Table Box	31
8.5.2	Sample Description Box	32
8.5.3	Degradation Priority Box	34
8.5.4	Sample Scale Box	34
8.6	Track Time Structures	34
8.6.1	Time to Sample Boxes	34
8.6.2	Sync Sample Box	39
8.6.3	Shadow Sync Sample Box	40
8.6.4	Independent and Disposable Samples Box	41
8.6.5	Edit Box	43
8.6.6	Edit List Box	43
8.7	Track Data Layout Structures	45
8.7.1	Data Information Box	45
8.7.2	Data Reference Box	45
8.7.3	Sample Size Boxes	46
8.7.4	Sample To Chunk Box	48
8.7.5	Chunk Offset Box	48
8.7.6	Padding Bits Box	49
8.7.7	Sub-Sample Information Box	50
8.7.8	Sample Auxiliary Information Sizes Box	51
8.7.9	Sample Auxiliary Information Offsets Box	53
8.8	Movie Fragments	54
8.8.1	Movie Extends Box	54
8.8.2	Movie Extends Header Box	54
8.8.3	Track Extends Box	55
8.8.4	Movie Fragment Box	56
8.8.5	Movie Fragment Header Box	56
8.8.6	Track Fragment Box	57
8.8.7	Track Fragment Header Box	57
8.8.8	Track Fragment Run Box	58
8.8.9	Movie Fragment Random Access Box	59
8.8.10	Track Fragment Random Access Box	60
8.8.11	Movie Fragment Random Access Offset Box	61
8.8.12	Track fragment decode time	62
8.8.13	Level Assignment Box	63
8.8.14	Sample Auxiliary Information in Movie Fragments	65
8.8.15	Track Extension Properties Box	65
8.8.16	Alternative Startup Sequence Properties Box	66
8.8.17	Metadata and user data in movie fragments	66
8.9	Sample Group Structures	67
8.9.1	Introduction	67
8.9.2	Sample to Group Box	68

8.9.3	Sample Group Description Box.....	69
8.9.4	Representation of group structures in Movie Fragments.....	70
8.10	User Data	71
8.10.1	User Data Box	71
8.10.2	Copyright Box	72
8.10.3	Track Selection Box.....	72
8.10.4	Track kind	74
8.11	Metadata Support.....	75
8.11.1	The Meta box	75
8.11.2	XML Boxes.....	76
8.11.3	The Item Location Box	76
8.11.4	Primary Item Box.....	79
8.11.5	Item Protection Box	80
8.11.6	Item Information Box.....	80
8.11.7	Additional Metadata Container Box.....	83
8.11.8	Metabox Relation Box	84
8.11.9	URL Forms for meta boxes.....	84
8.11.10	Static Metadata.....	85
8.11.11	Item Data Box	86
8.11.12	Item Reference Box.....	86
8.11.13	Auxiliary video metadata.....	87
8.12	Support for Protected Streams.....	87
8.12.1	Protection Scheme Information Box.....	89
8.12.2	Original Format Box.....	89
8.12.3	IPMPInfoBox.....	90
8.12.4	IPMP Control Box	90
8.12.5	Scheme Type Box	90
8.12.6	Scheme Information Box.....	91
8.13	File Delivery Format Support.....	91
8.13.1	Introduction.....	91
8.13.2	FD Item Information Box.....	91
8.13.3	File Partition Box	92
8.13.4	FEC Reservoir Box.....	94
8.13.5	FD Session Group Box.....	94
8.13.6	Group ID to Name Box	95
8.13.7	File Reservoir Box.....	96
8.14	Sub tracks.....	97
8.14.1	Introduction.....	97
8.14.2	Backward compatibility	97
8.14.3	Sub Track box.....	97
8.14.4	Sub Track Information box	98
8.14.5	Sub Track Definition box.....	99
8.14.6	Sub Track Sample Group box.....	100
8.15	Post-decoder requirements on media	100
8.15.1	General	100
8.15.2	Transformation	100

8.15.3	Restricted Scheme Information box.....	101
8.15.4	Scheme for stereoscopic video arrangements.....	102
8.16	Segments.....	104
8.16.1	Introduction.....	104
8.16.2	Segment Type Box.....	104
8.16.3	Segment Index Box.....	104
8.16.4	Subsegment Index Box.....	108
8.16.5	Producer Reference Time Box.....	110
8.17	Support for Incomplete Tracks.....	111
8.17.1	General.....	111
8.17.2	Transformation.....	112
8.17.3	Complete Track Information Box.....	113
9	Hint Track Formats.....	113
9.1	RTP and SRTP Hint Track Format.....	113
9.1.1	Introduction.....	113
9.1.2	Sample Description Format.....	114
9.1.3	Sample Format.....	115
9.1.4	SDP Information.....	118
9.1.5	Statistical Information.....	119
9.2	ALC/LCT and FLUTE Hint Track Format.....	120
9.2.1	Introduction.....	120
9.2.2	Design principles.....	121
9.2.3	Sample Description Format.....	122
9.2.4	Sample Format.....	123
9.3	MPEG-2 Transport Hint Track Format.....	126
9.3.1	Introduction.....	126
9.3.2	Design Principles.....	127
9.3.3	Sample Description Format.....	129
9.3.4	Sample Format.....	131
9.3.5	Protected MPEG 2 Transport Stream Hint Track.....	133
9.4	RTP, RTCP, SRTP and SRTCP Reception Hint Tracks.....	133
9.4.1	RTP Reception Hint Track.....	133
9.4.2	RTCP Reception Hint Track.....	137
9.4.3	SRTP Reception Hint Track.....	139
9.4.4	SRTCP Reception Hint Tracks.....	141
9.4.5	Protected RTP Reception Hint Track.....	142
9.4.6	Recording Procedure.....	142
9.4.7	Parsing Procedure.....	142
10	Sample Groups.....	142
10.1	Random Access Recovery Points.....	142
10.2	Rate Share Groups.....	143
10.2.1	Introduction.....	143
10.2.2	Rate Share Sample Group Entry.....	145
10.2.3	Relationship between tracks.....	146
10.2.4	Bitrate allocation.....	146

10.3	Alternative Startup Sequences	147
10.3.4	Examples	148
10.4	Random Access Point (RAP) Sample Grouping	150
10.5	Temporal level sample grouping	150
10.6	Stream access point sample group	151
11	Extensibility	152
11.1	Objects	152
11.2	Storage formats	153
11.3	Derived File formats	153
12	Media-specific definitions	154
12.1	Video media	154
12.1.1	Media handler	154
12.1.2	Video media header	154
12.1.3	Sample entry	154
12.1.4	Pixel Aspect Ratio and Clean Aperture	155
12.1.5	Colour information	157
12.2	Audio media	158
12.2.1	Media handler	158
12.2.2	Sound media header	158
12.2.3	Sample entry	158
12.2.4	Channel layout	160
12.2.5	Downmix Instructions	161
12.2.6	DRC Information	163
12.2.7	Audio stream loudness	164
12.3	Metadata media	166
12.3.1	Media handler	166
12.3.2	Media header	166
12.3.3	Sample entry	166
12.4	Hint media	168
12.4.1	Media handler	168
12.4.2	Hint media header	168
12.4.3	Sample entry	168
12.5	Text media	169
12.5.1	Media handler	169
12.5.2	Media header	169
12.5.3	Sample entry	169
12.6	Subtitle media	170
12.6.1	Media handler	170
12.6.2	Subtitle media header	170
12.6.3	Sample entry	170
12.7	Font media	171
12.7.1	Media handler	171
12.7.2	Media header	171
12.7.3	Sample entry	171
12.8	Transformed media	171

Annex A (informative) Overview and Introduction	172
A.1 Section Overview	172
A.2 Core Concepts	172
A.3 Physical structure of the media	173
A.4 Temporal structure of the media	173
A.5 Interleave	173
A.6 Composition	174
A.7 Random access	174
A.8 Fragmented movie files	175
Annex B (void)	177
Annex C (informative) Guidelines on deriving from this specification	178
C.1 Introduction	178
C.2 General Principles	178
C.2.1 General	178
C.2.2 Base layer operations	179
C.3 Boxes	179
C.4 Brand Identifiers	180
C.4.1 Introduction	180
C.4.2 Usage of the Brand	180
C.4.3 Introduction of a new brand	181
C.4.4 Player Guideline	181
C.4.5 Authoring Guideline	181
C.4.6 Example	182
C.5 Storage of new media types	182
C.6 Use of Template fields	182
C.7 Tracks	183
C.7.1 Data Location	183
C.7.2 Time	183
C.7.3 Media Types	184
C.7.4 Coding Types	184
C.7.5 Sub-sample information	184
C.7.6 Sample Dependency	184
C.7.7 Sample Groups	184
C.7.8 Track-level	185
C.7.9 Protection	185
C.8 Construction of fragmented movies	185
C.9 Meta-data	186
C.10 Registration	186
C.11 Guidelines on the use of sample groups, timed metadata tracks, and sample auxiliary information	186
Annex D (informative) Registration Authority	189
D.1 Code points to be registered	189
D.2 Procedure for the request of an MPEG-4 registered identifier value	190
D.3 Responsibilities of the Registration Authority	190
D.4 Contact information for the Registration Authority	190

D.5	Responsibilities of Parties Requesting a RID	191
D.6	Appeal Procedure for Denied Applications	191
D.7	Registration Application Form	191
D.7.1	Contact Information of organization requesting a RID	191
D.7.2	Request for a specific RID	192
D.7.3	Short description of RID that is in use and date system was implemented.....	192
D.7.4	Statement of an intention to apply the assigned RID.....	192
D.7.5	Date of intended implementation of the RID	192
D.7.6	Authorized representative	192
D.7.7	For official use of the Registration Authority	193
Annex E	(normative) File format brands	194
E.1	Introduction.....	194
E.2	The 'isom' brand.....	195
E.3	The 'avc1' brand.....	196
E.4	The 'iso2' brand.....	196
E.5	The 'mp71' brand.....	197
E.6	The 'iso3' brand.....	197
E.7	The 'iso4' brand.....	198
E.8	The 'iso5' brand.....	198
E.9	The 'iso6' brand.....	199
E.10	The 'iso7' brand	199
E.11	The 'iso8' brand	200
E.12	The 'iso9' brand	200
Annex F	(void).....	201
Annex G	(informative) URI-labelled metadata forms.....	202
G.1	UUID-labelled metadata	202
G.2	ISO OID-labelled metadata	202
G.3	SMPTE-labelled metadata.....	203
Annex H	(informative) Processing of RTP streams and reception hint tracks.....	204
H.1	Introduction	204
H.1.1	Overview.....	204
H.1.2	Structure	204
H.1.3	Terms and definitions.....	204
H.2	Synchronization of RTP streams.....	204
H.3	Recording of RTP streams.....	205
H.3.1	Introduction.....	205
H.3.2	Compensation for unequal starting for position of received RTP streams	208
H.3.3	Recording of SDP.....	209
H.3.4	Creation of a sample within an RTP reception hint track.....	209
H.3.5	Representation of RTP timestamps	210
H.3.6	Recording operations to facilitate inter-stream synchronization in playback.....	213
H.3.7	Representation of reception times.....	215
H.3.8	Creation of media samples	215
H.3.9	Creation of hint samples referring to media samples	216

H.4	Playing of recorded RTP streams	216
H.4.1	Introduction	216
H.4.2	Preparation for the playback	217
H.4.3	Decoding of a sample within an RTP reception hint track	217
H.4.4	Lip synchronization	218
H.4.5	Random access	219
H.5	Re-sending recorded RTP streams	220
H.5.1	Introduction	220
H.5.2	Re-sending RTP packets	220
H.5.3	RTCP Processing	222
Annex I (normative)	Stream Access Points	223
I.1	Introduction	223
I.2	SAP properties	223
I.2.1	General	223
I.2.2	SAP properties for layers	224
I.3	SAP types	225
Annex J (normative)	MIME Type Registration of Segments	226
J.1	Introduction	226
J.2	Registration	226
Annex K : Segment Index Examples (informative)		227
K.1	Introduction	227
K.2	Examples	227
K.2.1	Simple one-level indexing	227
K.2.2	Hierarchical	227
K.2.3	Daisy-chain	228
K.2.4	Combination hierarchical and daisy-chain	229

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. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

~~International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.~~

~~The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.~~

~~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).~~

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).

~~ISO/IEC 14496-12 was prepared by Joint Technical Committee ISO/IEC JTC 1, Information technology, Subcommittee 1.~~
Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: [Foreword - Supplementary information](#)

The committee responsible for this document is ISO/IEC JTC 1, Information technology, SC 29, Coding of audio, picture, multimedia and hypermedia information.

This fifth edition cancels and replaces the fourth edition (~~ISO/IEC 14496-12:2012~~), which has been technically revised. It also incorporates the Amendments ISO/IEC 14496-12:2012/Amd1:2013, ISO/IEC 14496-12:2012/Amd2:2014, ISO/IEC 14496-12:2012/Amd3:2015 and the Technical Corrigenda ISO/IEC 14496-12:2012/Cor1:2013, ISO/IEC 14496-12:2012/Cor2:2014 and ISO/IEC 14496-12:2012/Cor3:2015.

ISO/IEC 14496-12:2015(E)

ISO/IEC 14496 consists of the following parts, under the general title *Information technology — Coding of audio-visual objects*:

- *Part 1: Systems*
- *Part 2: Visual*
- *Part 3: Audio*
- *Part 4: Conformance testing*
- *Part 5: Reference software*
- *Part 6: Delivery Multimedia Integration Framework (DMIF)*
- *Part 7: Optimized reference software for coding of audio-visual objects*
- *Part 8: Carriage of ISO/IEC 14496 contents over IP networks*
- *Part 9: Reference hardware description*
- *Part 10: Advanced Video Coding*
- *Part 11: Scene description and application engine*
- *Part 12: ISO base media file format*
- *Part 13: Intellectual Property Management and Protection (IPMP) extensions*
- *Part 14: MP4 file format*
- *Part 15: Carriage of NAL unit structured video in the ISO Base Media File Format*
- *Part 16: Animation Framework eXtension (AFX)*
- *Part 17: Streaming text format*
- *Part 18: Font compression and streaming*
- *Part 19: Synthesized texture stream*
- *Part 20: Lightweight Application Scene Representation (LSeR) and Simple Aggregation Format (SAF)*
- *Part 21: MPEG-J Graphics Framework eXtensions (GFX)*
- *Part 22: Open Font Format*
- *Part 23: Symbolic Music Representation*
- *Part 24: Audio and systems interaction*
- *Part 25: 3D Graphics Compression Model*

- *Part 26: Audio conformance*
- *Part 27: 3D Graphics conformance*
- *Part 28: Composite font representation*
- *Part 29: Web video coding*
- *Part 30: Timed text and other visual overlays in ISO base media file format*
- *Part 31: Video Coding for Browsers*

Introduction

The ISO Base Media File Format is designed to contain timed media information for a presentation in a flexible, extensible format that facilitates interchange, management, editing, and presentation of the media. This presentation may be 'local' to the system containing the presentation, or may be via a network or other stream delivery mechanism.

The file structure is object-oriented; a file can be decomposed into constituent objects very simply, and the structure of the objects inferred directly from their type.

The file format is designed to be independent of any particular network protocol while enabling efficient support for them in general.

The ISO Base Media File Format is a base format for media file formats.

It is intended that the ISO Base Media File Format shall be jointly maintained by WG1 and WG11. Consequently, a subdivision of work created ISO/IEC 15444-12 and ISO/IEC 14496-12 in order to document the ISO Base Media File Format and to facilitate the joint maintenance.

This technically identical text is published as ISO/IEC 14496-12 for MPEG-4, and as ISO/IEC 15444-12 for JPEG 2000, and reference to this specification should be made accordingly. The recommendation is to reference one, for example ISO/IEC 14496-12, and append to the reference a parenthetical comment identifying the other, for example "(technically identical to ISO/IEC 15444-12)".

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 patents.

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

The holder of this patent right has assured the ISO and IEC that he 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 the ISO and IEC. Information may be obtained from the companies listed in Annex B.

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

ISO (www.iso.org/patents) and IEC (<http://patents.iec.ch>) maintain on-line databases of patents relevant to their standards. Users are encouraged to consult the databases for the most up to date information concerning patents.

Information technology — Coding of audio-visual objects —

Part 12:

ISO base media file format

1 Scope

This part of ISO/IEC 14496 specifies the ISO base media file format, which is a general format forming the basis for a number of other more specific file formats. This format contains the timing, structure, and media information for timed sequences of media data, such as audio-visual presentations.

This part of ISO/IEC 14496 is applicable to MPEG-4, but its technical content is identical to that of ISO/IEC 15444-12, which is applicable to JPEG 2000.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 639-2:1998, *Codes for the representation of names of languages — Part 2: Alpha-3 code*

ISO/IEC 9834-8:2005, *Information technology — Open Systems Interconnection — Procedures for the operation of OSI Registration Authorities: Generation and registration of Universally Unique Identifiers (UUIDs) and their use as ASN.1 Object Identifier components*

ISO/IEC 11578:1996, *Information technology — Open Systems Interconnection — Remote Procedure Call (RPC)*

ISO/IEC 14496-1:2010, *Information technology — Coding of audio-visual objects — Part 1: Systems*

ISO/IEC 14496-10, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

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

ISO/IEC 15444-1, *Information technology — JPEG 2000 image coding system: Core coding system*

ISO/IEC 15444-3, *Information technology — JPEG 2000 image coding system: Motion JPEG 2000*

ISO/IEC 15938-1, *Information technology — Multimedia content description interface — Part 1: Systems*

ISO/IEC 23001-1, *Information technology — MPEG systems technologies — Part 1: Binary MPEG format for XML*

ISO/IEC 14496-12:2015(E)

ISO/IEC 23002-3, *Information technology — MPEG video technologies — Part 3: Representation of auxiliary video and supplemental information*

ISO/IEC 29199-2:2012, *Information technology — JPEG XR image coding system — Part 2: Image coding specification*

ISO 15076-1:2010, *Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010*

IETF RFC 2045, *Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies*, FREED, N. and BORENSTEIN, N., November 1996

IETF RFC 2046, *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*, FREED, N. and BORENSTEIN, N., November 1996

IETF RFC 3550, *RTP: A Transport Protocol for Real-Time Applications*, SCHULZRINNE, H. et al., July 2003.

IETF RFC 3711, *The Secure Real-time Transport Protocol (SRTP)*, BAUGHER, M. et al., March 2004

IETF RFC 5052, *Forward Error Correction (FEC) Building Block*, WATSON, M. et al., August 2007

IETF RFC 5905, *Network Time Protocol Version 4: Protocol and Algorithms Specification*, MILLS, D., et al, June 2010

SMIL 1.0, *Synchronized Multimedia Integration Language (SMIL) 1.0 Specification*, <<http://www.w3.org/TR/REC-smil/>>

Rec. ITU-R TF.460-6, *Standard-frequency and time-signal emissions (Annex I for the definition of UTC.)*

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

ITU-R, Recommendation ITU-R BS.1770-3. *Algorithm to measure audio programme loudness and true-peak audio level*, August 2012.

ITU-R, Recommendation ITU-R BS.1771-1. *Requirements for loudness and true-peak indicating meters*, January 2012.

EBU R 128-2014, *Loudness normalization and permitted maximum level of audio signals*, June 2014.

EBU EBU – Tech 3341, *Loudness Metering: EBU mode metering to supplement loudness normalization in accordance with EBU R128*

EBU EBU-Tech 3342, *Loudness Range: A measure to supplement loudness normalisation in accordance with EBU R 128*, Geneva, August 2011

ETSI TS 101 154 V1.11.1, *Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*, November 2012.

ATSC Document A/85:2011, *ATSC Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television*, July 2011

ATSC Doc. A/52:2012, *ATSC Standard: Digital Audio Compression (AC-3, E-AC-3)*.

IETF RFC 5646, BCP 47, *Tags for Identifying Languages*, PHILLIPS, A., et al, September 2009

3 Terms, definitions, and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1.1

box

object-oriented building block defined by a unique type identifier and length

NOTE—[Note 1 to entry](#): Called 'atom' in some specifications, including the first definition of MP4.

3.1.2

chunk

contiguous set of samples for one track

3.1.3

container box

box whose sole purpose is to contain and group a set of related boxes

NOTE—[Note 1 to entry](#): Container boxes are normally not derived from 'fullbox'.

3.1.4

hint track

special track which does not contain media data, but instead contains instructions for packaging one or more tracks into a streaming channel

3.1.5

hinter

tool that is run on a file containing only media, to add one or more hint tracks to the file and so facilitate streaming

3.1.6

ISO Base Media File

name of the files conforming to the file format described in this specification

3.1.7

leaf subsegment

subsegment that does not contain any indexing information that would enable its further division into subsegments