

Fourth edition
2012-07-15

Corrected version
2012-09-15

**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 14496-12:2012(E)



© ISO/IEC 2012



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2012

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from 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

Published in Switzerland

Contents

Page

Foreword	ix
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	4
4 Object-structured File Organization	4
4.1 File Structure	4
4.2 Object Structure	4
4.3 File Type Box	5
4.3.1 Definition	5
5 Design Considerations	6
5.1 Usage	6
5.1.1 Introduction.....	6
5.1.2 Interchange	6
5.1.3 Content Creation	8
5.1.4 Preparation for streaming	8
5.1.5 Local presentation.....	8
5.1.6 Streamed presentation.....	9
5.2 Design principles.....	9
6 ISO Base Media File organization.....	10
6.1 Presentation structure	10
6.1.1 File Structure	10
6.1.2 Object Structure	10
6.1.3 Meta Data and Media Data	10
6.1.4 Track Identifiers	10
6.2 Metadata Structure (Objects)	10
6.2.1 Box	10
6.2.2 Data Types and fields.....	11
6.2.3 Box Order	12
6.2.4 URIs as type indicators.....	14
6.3 Brand Identification.....	14
7 Streaming Support	15
7.1 Handling of Streaming Protocols	15
7.2 Protocol 'hint' tracks.....	15
7.3 Hint Track Format.....	16
8 Box Structures	17
8.1 File Structure and general boxes.....	17
8.1.1 Media Data Box.....	17
8.1.2 Free Space Box.....	17
8.1.3 Progressive Download Information Box	18
8.2 Movie Structure	18
8.2.1 Movie Box.....	18
8.2.2 Movie Header Box	19
8.3 Track Structure	20
8.3.1 Track Box	20
8.3.2 Track Header Box	20

8.3.3	Track Reference Box	22
8.3.4	Track Group Box	23
8.4	Track Media Structure	24
8.4.1	Media Box	24
8.4.2	Media Header Box	24
8.4.3	Handler Reference Box	25
8.4.4	Media Information Box	26
8.4.5	Media Information Header Boxes	26
8.5	Sample Tables	28
8.5.1	Sample Table Box	28
8.5.2	Sample Description Box	28
8.5.3	Degradation Priority Box	34
8.5.4	Sample Scale Box	35
8.6	Track Time Structures	35
8.6.1	Time to Sample Boxes	35
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	42
8.6.6	Edit List Box	42
8.7	Track Data Layout Structures	44
8.7.1	Data Information Box	44
8.7.2	Data Reference Box	44
8.7.3	Sample Size Boxes	45
8.7.4	Sample To Chunk Box	46
8.7.5	Chunk Offset Box	47
8.7.6	Padding Bits Box	48
8.7.7	Sub-Sample Information Box	49
8.7.8	Sample Auxiliary Information Sizes Box	50
8.7.9	Sample Auxiliary Information Offsets Box	51
8.8	Movie Fragments	52
8.8.1	Movie Extends Box	52
8.8.2	Movie Extends Header Box	53
8.8.3	Track Extends Box	53
8.8.4	Movie Fragment Box	54
8.8.5	Movie Fragment Header Box	54
8.8.6	Track Fragment Box	55
8.8.7	Track Fragment Header Box	55
8.8.8	Track Fragment Run Box	56
8.8.9	Movie Fragment Random Access Box	57
8.8.10	Track Fragment Random Access Box	58
8.8.11	Movie Fragment Random Access Offset Box	59
8.8.12	Track fragment decode time	59
8.8.13	Level Assignment Box	60
8.8.14	Sample Auxiliary Information in Movie Fragments	62
8.9	Sample Group Structures	62
8.9.1	Introduction	62
8.9.2	Sample to Group Box	63
8.9.3	Sample Group Description Box	64
8.9.4	Representation of group structures in Movie Fragments	65
8.10	User Data	66
8.10.1	User Data Box	66
8.10.2	Copyright Box	66
8.10.3	Track Selection Box	67
8.11	Metadata Support	69
8.11.1	The Meta box	69
8.11.2	XML Boxes	70
8.11.3	The Item Location Box	70
8.11.4	Primary Item Box	72
8.11.5	Item Protection Box	73

8.11.6	Item Information Box.....	73
8.11.7	Additional Metadata Container Box.....	75
8.11.8	Metabox Relation Box.....	76
8.11.9	URL Forms for meta boxes	76
8.11.10	Static Metadata	77
8.11.11	Item Data Box.....	78
8.11.12	Item Reference Box	78
8.11.13	Auxiliary video metadata	79
8.12	Support for Protected Streams	79
8.12.1	Protection Scheme Information Box	80
8.12.2	Original Format Box	81
8.12.3	IPMPInfoBox	81
8.12.4	IPMP Control Box	81
8.12.5	Scheme Type Box.....	81
8.12.6	Scheme Information Box	82
8.13	File Delivery Format Support	82
8.13.1	Introduction.....	82
8.13.2	FD Item Information Box.....	83
8.13.3	File Partition Box	83
8.13.4	FEC Reservoir Box	85
8.13.5	FD Session Group Box	85
8.13.6	Group ID to Name Box	86
8.13.7	File Reservoir Box	87
8.14	Sub tracks	87
8.14.1	Introduction.....	87
8.14.2	Backward compatibility	88
8.14.3	Sub Track box.....	88
8.14.4	Sub Track Information box.....	88
8.14.5	Sub Track Definition box	89
8.14.6	Sub Track Sample Group box	90
8.15	Post-decoder requirements on media.....	90
8.15.1	General	90
8.15.2	Transformation	90
8.15.3	Restricted Scheme Information box.....	91
8.15.4	Scheme for stereoscopic video arrangements	91
8.16	Segments	93
8.16.1	Introduction.....	93
8.16.2	Segment Type Box	93
8.16.3	Segment Index Box	94
8.16.4	Subsegment Index Box.....	97
8.16.5	Producer Reference Time Box	99
9	Hint Track Formats.....	100
9.1	RTP and SRTP Hint Track Format	100
9.1.1	Introduction.....	100
9.1.2	Sample Description Format.....	100
9.1.3	Sample Format.....	102
9.1.4	SDP Information	105
9.1.5	Statistical Information.....	105
9.2	ALC/LCT and FLUTE Hint Track Format	106
9.2.1	Introduction.....	106
9.2.2	Design principles.....	107
9.2.3	Sample Description Format.....	108
9.2.4	Sample Format.....	109
9.3	MPEG-2 Transport Hint Track Format	112
9.3.1	Introduction.....	112
9.3.2	Design Principles	112
9.3.3	Sample Description Format.....	114
9.3.4	Sample Format.....	116
9.3.5	Protected MPEG 2 Transport Stream Hint Track	118

9.4	RTP, RTCP, SRTP and SRTCP Reception Hint Tracks	118
9.4.1	RTP Reception Hint Track.....	118
9.4.2	RTCP Reception Hint Track	122
9.4.3	SRTP Reception Hint Track	123
9.4.4	SRTCP Reception Hint Tracks.....	125
9.4.5	Protected RTP Reception Hint Track	126
9.4.6	Recording Procedure	126
9.4.7	Parsing Procedure	126
10	Sample Groups	126
10.1	Random Access Recovery Points.....	126
10.2	Rate Share Groups	127
10.2.1	Introduction	127
10.2.2	Rate Share Sample Group Entry	128
10.2.3	Relationship between tracks	129
10.2.4	Bitrate allocation.....	130
10.3	Alternative Startup Sequences.....	130
10.3.1	Definition	130
10.3.2	Syntax	131
10.3.3	Semantics	131
10.3.4	Examples	131
10.4	Random Access Point (RAP) Sample Grouping.....	133
10.4.1	Definition	133
10.4.2	Syntax	133
10.4.3	Semantics	133
10.5	Temporal level sample grouping.....	133
10.5.1	Definition	133
10.5.2	Syntax	134
10.5.3	Semantics	134
11	Extensibility.....	134
11.1	Objects.....	134
11.2	Storage formats	135
11.3	Derived File formats	135
Annex A	(informative) Overview and Introduction.....	136
A.1	Section Overview	136
A.2	Core Concepts	136
A.3	Physical structure of the media	136
A.4	Temporal structure of the media.....	137
A.5	Interleave	137
A.6	Composition.....	137
A.7	Random access.....	138
A.8	Fragmented movie files.....	138
Annex B	(informative) Patent Statements.....	140
Annex C	(informative) Guidelines on deriving from this specification.....	141
C.1	Introduction	141
C.2	General Principles	141
C.2.1	General.....	141
C.2.2	Base layer operations	141
C.3	Boxes	142
C.4	Brand Identifiers	142
C.4.1	Introduction	142
C.4.2	Usage of the Brand	143
C.4.3	Introduction of a new brand	143
C.4.4	Player Guideline.....	143
C.4.5	Authoring Guideline	144
C.4.6	Example	144
C.5	Storage of new media types	144
C.6	Use of Template fields.....	145

C.7	Tracks	145
C.7.1	Data Location	145
C.7.2	Time	145
C.7.3	Media Types	146
C.7.4	Coding Types	146
C.7.5	Sub-sample information	146
C.7.6	Sample Dependency	146
C.7.7	Sample Groups	146
C.7.8	Track-level	146
C.7.9	Protection	147
C.8	Construction of fragmented movies	147
C.9	Meta-data	148
C.10	Registration	148
C.11	Guidelines on the use of sample groups, timed metadata tracks, and sample auxiliary information	148
Annex D	(informative) Registration Authority	150
D.1	Code points to be registered	150
D.2	Procedure for the request of an MPEG-4 registered identifier value	150
D.3	Responsibilities of the Registration Authority	151
D.4	Contact information for the Registration Authority	151
D.5	Responsibilities of Parties Requesting a RID	151
D.6	Appeal Procedure for Denied Applications	152
D.7	Registration Application Form	152
D.7.1	Contact Information of organization requesting a RID	152
D.7.2	Request for a specific RID	152
D.7.3	Short description of RID that is in use and date system was implemented	153
D.7.4	Statement of an intention to apply the assigned RID	153
D.7.5	Date of intended implementation of the RID	153
D.7.6	Authorized representative	153
D.7.7	For official use of the Registration Authority	153
Annex E	(normative) File format brands	154
E.1	Introduction	154
E.2	The 'isom' brand	155
E.3	The 'avc1' brand	156
E.4	The 'iso2' brand	156
E.5	The 'mp71' brand	157
E.6	The 'iso3' brand	157
E.7	The 'iso4' brand	157
E.8	The 'iso5' brand	158
E.9	The 'iso6' brand	158
Annex F	(informative) Document Cross-Reference	159
Annex G	(informative) URI-labelled metadata forms	161
G.1	UUID-labelled metadata	161
G.2	ISO OID-labelled metadata	161
G.3	SMPTE-labelled metadata	161
Annex H	(informative) Processing of RTP streams and reception hint tracks	163
H.1	Introduction	163
H.1.1	Overview	163
H.1.2	Structure	163
H.1.3	Terms and definitions	163
H.2	Synchronization of RTP streams	163
H.3	Recording of RTP streams	164
H.3.1	Introduction	164
H.3.2	Compensation for unequal starting for position of received RTP streams	166
H.3.3	Recording of SDP	167
H.3.4	Creation of a sample within an RTP reception hint track	167
H.3.5	Representation of RTP timestamps	168

H.3.6	Recording operations to facilitate inter-stream synchronization in playback	171
H.3.7	Representation of reception times.....	172
H.3.8	Creation of media samples	173
H.3.9	Creation of hint samples referring to media samples.....	173
H.4	Playing of recorded RTP streams	173
H.4.1	Introduction	173
H.4.2	Preparation for the playback	174
H.4.3	Decoding of a sample within an RTP reception hint track	174
H.4.4	Lip synchronization	174
H.4.5	Random access.....	176
H.5	Re-sending recorded RTP streams.....	176
H.5.1	Introduction	176
H.5.2	Re-sending RTP packets.....	177
H.5.3	RTCP Processing.....	178
Annex I	(normative) Stream Access Points	179
I.1	Introduction	179
I.2	SAP properties	179
I.3	SAP types	179
Annex J	(normative) MIME Type Registration of Segments	181
J.1	Introduction	181
J.2	Registration	181
Bibliography	182

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.

ISO/IEC 14496-12 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This fourth edition cancels and replaces the third edition (ISO/IEC 14496-12:2008) of which it constitutes a minor revision. It also incorporates the Amendment ISO/IEC 14496-12:2008/Amd.1:2009 and the Technical Corrigenda ISO/IEC 14496-12:2008/Cor.1:2008, ISO/IEC 14496-12:2008/Cor.2:2009, ISO/IEC 14496-12:2008/Cor.3:2009, and ISO/IEC 14496-12:2008/Cor.4:2011.

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* [Technical Report]
- *Part 8: Carriage of ISO/IEC 14496 contents over IP networks*
- *Part 9: Reference hardware description* [Technical Report]
- *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: Advanced Video Coding (AVC) 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 (LASER) 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 [Technical Report]*
- *Part 25: 3D Graphics Compression Model*
- *Part 26: Audio conformance*
- *Part 27: 3D Graphics conformance*
- *Part 28: Composite font representation*

This corrected version of ISO/IEC 14496-12:2012 incorporates the corrections made by ISO/IEC 14496-12:2008 draft Technical Corrigendum 5 (unpublished).

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.

This is a free preview. Purchase the entire publication at the link below:

- ▶ Looking for additional Standards? Visit [SAI Global Infostore](#)
- ▶ Subscribe to our [Free Newsletters about Australian Standards® in Legislation; ISO, IEC, BSI and more](#)
- ▶ Do you need to [Manage Standards Collections Online?](#)
- ▶ Learn about [LexConnect, All Jurisdictions, Standards referenced in Australian legislation](#)
- ▶ Do you want to [know when a Standard has changed?](#)
- ▶ Want to [become an SAI Global Standards Sales Affiliate?](#)

Learn about other SAI Global Services:

- ▶ [LOGICOM Military Parts and Supplier Database](#)
- ▶ [Metals Infobase Database of Metal Grades, Standards and Manufacturers](#)
- ▶ [Materials Infobase Database of Materials, Standards and Suppliers](#)
- ▶ [Database of European Law, CELEX and Court Decisions](#)

Need to speak with a Customer Service Representative - [Contact Us](#)