

Second edition
2005-04-01

Corrected version
2005-10-01

**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 14496-12:2005(E)



PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

© ISO/IEC 2005

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.....	vi
Introduction	viii
1 Scope	1
2 Normative references	1
3 Definitions	1
4 Object-structured File Organization	3
4.1 File Structure	3
4.2 Object Structure	3
4.3 File Type Box	4
4.3.1 Definition	4
4.3.2 Syntax	5
4.3.3 Semantics	5
5 Design Considerations	5
5.1 Usage	5
5.1.1 Introduction	5
5.1.2 Interchange	5
5.1.3 Content Creation	6
5.1.4 Preparation for streaming	6
5.1.5 Local presentation	7
5.1.6 Streamed presentation	7
5.2 Design principles	7
6 ISO Base Media File organization	8
6.1 Presentation structure	8
6.1.1 File Structure	8
6.1.2 Object Structure	8
6.1.3 Meta Data and Media Data	8
6.1.4 Track Identifiers	8
6.2 Metadata Structure (Objects)	9
6.2.1 Box	9
6.2.2 Data Types and fields	9
6.2.3 Box Order	10
7 Streaming Support	13
7.1 Handling of Streaming Protocols	13
7.2 Protocol 'hint' tracks	13
7.3 Hint Track Format	14
8 Box Definitions	14
8.1 Movie Box	14
8.2 Media Data Box	15
8.3 Movie Header Box	15
8.4 Track Box	16
8.5 Track Header Box	17
8.6 Track Reference Box	18
8.7 Media Box	19
8.8 Media Header Box	19
8.9 Handler Reference Box	20
8.10 Media Information Box	21
8.11 Media Information Header Boxes	21
8.11.2 Video Media Header Box	21

8.11.3	Sound Media Header Box.....	22
8.11.4	Hint Media Header Box.....	22
8.11.5	Null Media Header Box.....	22
8.12	Data Information Box.....	23
8.13	Data Reference Box.....	23
8.14	Sample Table Box.....	24
8.15	Time to Sample Boxes.....	24
8.15.2	Decoding Time to Sample Box.....	25
8.15.3	Composition Time to Sample Box.....	26
8.16	Sample Description Box.....	27
8.17	Sample Size Boxes.....	29
8.17.2	Sample Size Box.....	30
8.17.3	Compact Sample Size Box.....	30
8.18	Sample To Chunk Box.....	30
8.19	Chunk Offset Box.....	31
8.20	Sync Sample Box.....	32
8.21	Shadow Sync Sample Box.....	32
8.22	Degradation Priority Box.....	33
8.23	Padding Bits Box.....	34
8.24	Free Space Box.....	34
8.25	Edit Box.....	35
8.26	Edit List Box.....	35
8.27	User Data Box.....	36
8.28	Copyright Box.....	37
8.29	Movie Extends Box.....	37
8.30	Movie Extends Header Box.....	37
8.31	Track Extends Box.....	38
8.32	Movie Fragment Box.....	39
8.33	Movie Fragment Header Box.....	39
8.34	Track Fragment Box.....	40
8.35	Track Fragment Header Box.....	40
8.36	Track Fragment Run Box.....	41
8.37	Movie Fragment Random Access Box.....	42
8.38	Track Fragment Random Access Box.....	43
8.39	Movie Fragment Random Access Offset Box.....	44
8.40	AVC Extensions.....	44
8.40.2	Independent and Disposable Samples Box.....	44
8.40.3	Sample Groups.....	45
8.40.4	Random Access Recovery Points.....	48
8.41	Sample Scale Box.....	49
8.42	Sub-Sample Information Box.....	50
8.43	Progressive Download Information Box.....	51
8.44	Metadata Support.....	52
8.44.1	The Metadata Box.....	52
8.44.2	XML Boxes.....	53
8.44.3	The Item Location Box.....	53
8.44.4	Primary Item Box.....	55
8.44.5	Item Protection Box.....	55
8.44.6	Item Information Box.....	56
8.44.7	URL Forms for meta boxes.....	56
8.44.8	Static Metadata.....	57
8.45	Support for Protected Streams.....	58
8.45.1	Protection Scheme Information Box.....	59
8.45.2	Original Format Box.....	59
8.45.3	IPMPInfoBox.....	60
8.45.4	IPMP Control Box.....	60
8.45.5	Scheme Type Box.....	61
8.45.6	Scheme Information Box.....	62
9	Extensibility.....	62

9.1	Objects	62
9.2	Storage formats	63
9.3	Derived File formats	63
10	RTP and SRTP Hint Track Format.....	63
10.1	Introduction	63
10.2	Sample Description Format	64
10.2.1	SRTP Process box 'srpp':.....	65
10.3	Sample Format	66
10.3.1	Packet Entry format	66
10.3.2	Constructor format	67
10.4	SDP Information	68
10.4.1	Movie SDP information	68
10.4.2	Track SDP Information	68
10.5	Statistical Information	69
Annex A	(informative) Overview and Introduction.....	70
A.1	Section Overview	70
A.2	Core Concepts	70
A.3	Physical structure of the media	70
A.4	Temporal structure of the media.....	71
A.5	Interleave	71
A.6	Composition	71
A.7	Random access.....	72
A.8	Fragmented movie files.....	72
Annex B	(informative) Patent Statements	74
Annex C	(informative) Guidelines on deriving from this specification.....	75
C.1	Introduction	75
C.2	General Principles	75
C.3	Brand Identifiers	75
C.3.1	Introduction	75
C.3.2	Usage of the Brand	75
C.3.3	Introduction of a new brand	76
C.3.4	Player Guideline	76
C.3.5	Example	76
C.4	Box layout and order	77
C.5	Storage of new media types	77
C.6	Use of Template fields.....	77
C.7	Construction of fragmented movies	77
Annex D	(informative) Registration Authority	79
D.1	Code points to be registered	79
D.2	Procedure for the request of an MPEG-4 registered identifier value	79
D.3	Responsibilities of the Registration Authority	80
D.4	Contact information for the Registration Authority	80
D.5	Responsibilities of Parties Requesting a RID	80
D.6	Appeal Procedure for Denied Applications	81
D.7	Registration Application Form	81
D.7.1	Contact Information of organization requesting a RID	81
D.7.2	Request for a specific RID	81
D.7.3	Short description of RID that is in use and date system was implemented	82
D.7.4	Statement of an intention to apply the assigned RID	82
D.7.5	Date of intended implementation of the RID	82
D.7.6	Authorized representative	82
D.7.7	For official use of the Registration Authority	83
Bibliography	84

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.

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.

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 second edition cancels and replaces the first edition (ISO/IEC 14496-12:2004) which has been technically revised.

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: 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 (LSeR) and Simple Aggregation Format (SAF)*

The following parts are under preparation:

- *Part 21: MPEG-J extensions for rendering*

This corrected version of ISO/IEC 14496-12:2005 (E) incorporates the following correction:

- In 8.16.2, 8.17.2.1, 8.17.3.1, 8.18.2 and 8.19.2, “•” has been replaced with “≤”.

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 15444-12 and 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)".

Information technology — Coding of audio-visual objects —

Part 12:

ISO base media file format

1 Scope

This International Standard 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 referenced documents are indispensable for the application of this document. 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 11578:1996: *Information technology — Open Systems Interconnection — Remote Procedure Call (RPC)*

ISO/IEC 14496-1:2004: *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*

ITU-T Rec.T.800 | 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*

IETF RFC 3711, *"The Secure Real-time Transport Protocol"*, Baugher M. et al., March 2004.

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

3 Definitions

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

3.1

Box:

An object-oriented building block defined by a unique type identifier and length (called 'atom' in some specifications, including the first definition of MP4).

3.2

Chunk:

A contiguous set of samples for one track.

3.3

Container Box:

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

3.4

Hint Track:

A special track which does not contain media data. Instead it contains instructions for packaging one or more tracks into a streaming channel.

3.5

Hint:

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

3.6

Movie Box:

A container box whose sub-boxes define the metadata for a presentation ('moov').

3.7

Media Data Box:

A container box which can hold the actual media data for a presentation ('mdat').

3.8

ISO Base Media File:

The name of the file format described in this specification.

3.9

Presentation:

One or more motion sequences (q.v.), possibly combined with audio.

3.10

Sample:

In non-hint tracks, a sample is an individual frame of video, a time-contiguous series of video frames, or a time-contiguous compressed section of audio. In hint tracks, a sample defines the formation of one or more streaming packets. No two samples within a track may share the same time-stamp.

3.11

Sample Description:

A structure which defines and describes the format of some number of samples in a track.