

STANDARD

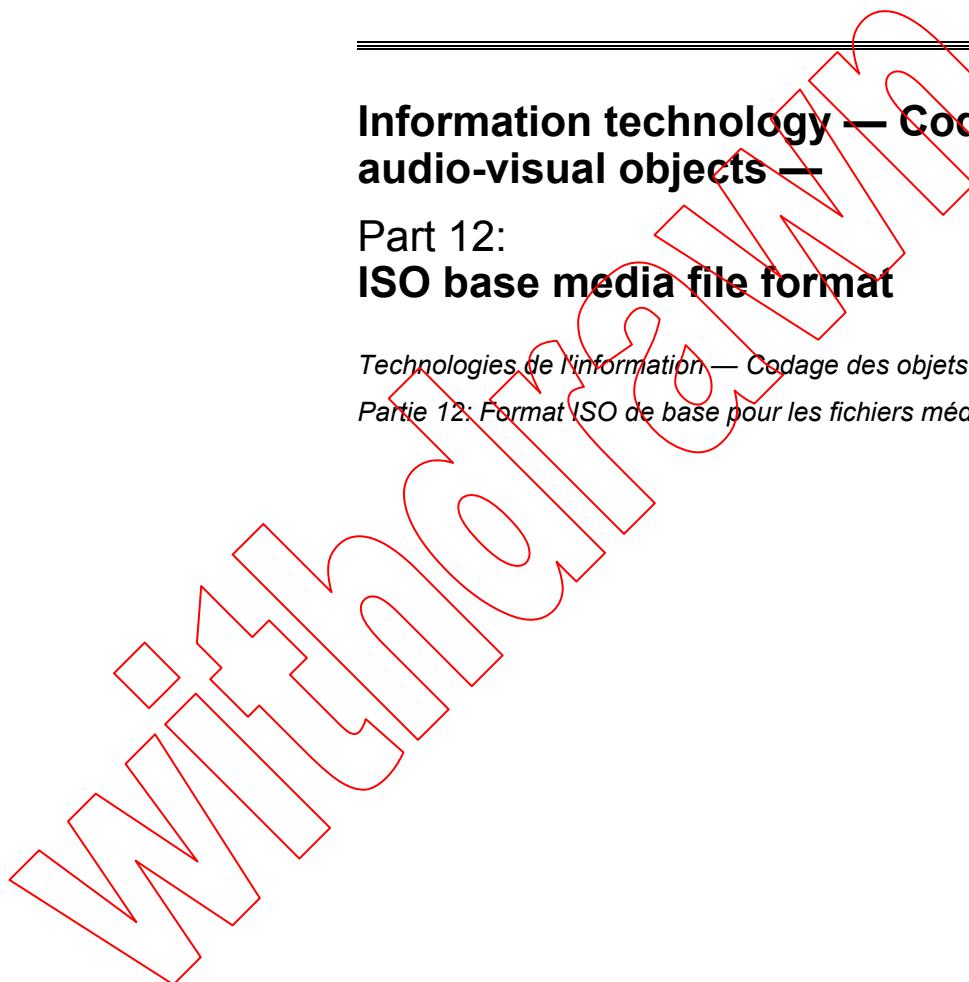
**ISO/IEC
14496-12**

Fourth edition
2012-07-15

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

Withdrawn

**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	7
5.1.4 Preparation for streaming	7
5.1.5 Local presentation.....	7
5.1.6 Streamed presentation.....	8
5.2 Design principles.....	8
6 ISO Base Media File organization	9
6.1 Presentation structure	9
6.1.1 File Structure	9
6.1.2 Object Structure	9
6.1.3 Meta Data and Media Data	9
6.1.4 Track Identifiers.....	9
6.2 Metadata Structure (Objects)	9
6.2.1 Box	9
6.2.2 Data Types and fields.....	10
6.2.3 Box Order	11
6.2.4 URIs as type indicators	13
6.3 Brand Identification.....	13
7 Streaming Support	14
7.1 Handling of Streaming Protocols	14
7.2 Protocol 'hint' tracks	14
7.3 Hint Track Format.....	15
8 Box Structures	16
8.1 File Structure and general boxes	16
8.1.1 Media Data Box	16
8.1.2 Free Space Box	16
8.1.3 Progressive Download Information Box	17
8.2 Movie Structure	17
8.2.1 Movie Box	17
8.2.2 Movie Header Box	18
8.3 Track Structure	19
8.3.1 Track Box	19
8.3.2 Track Header Box	19

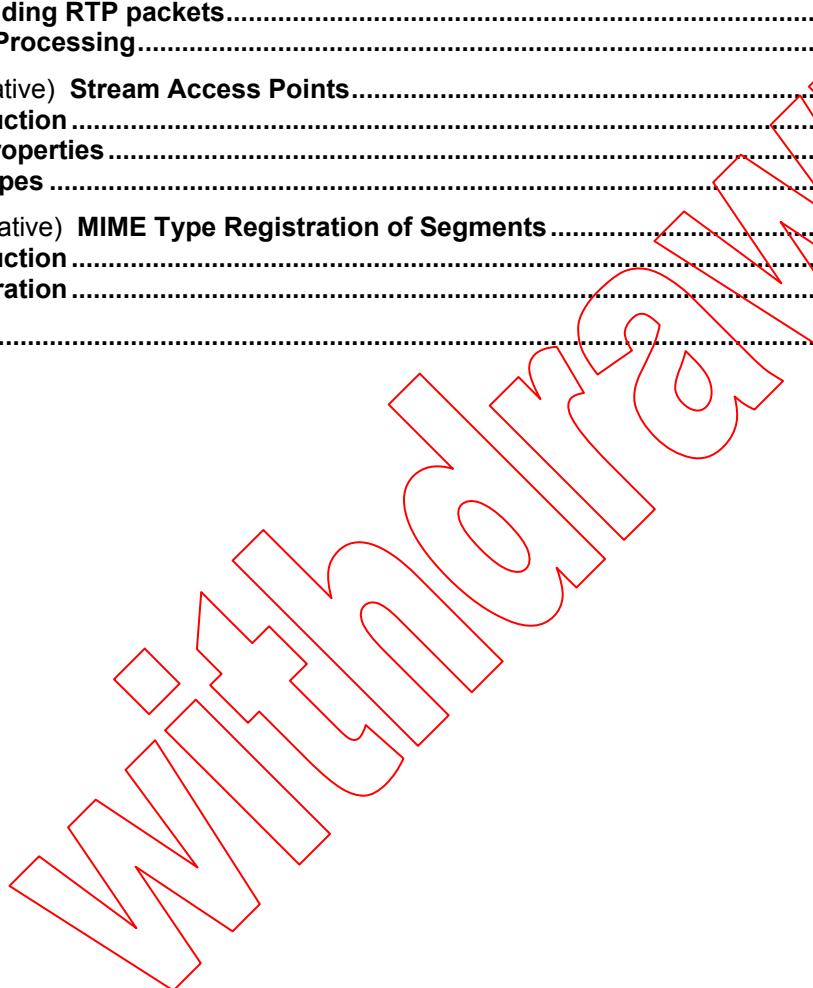
8.3.3	Track Reference Box	21
8.3.4	Track Group Box.....	22
8.4	Track Media Structure	23
8.4.1	Media Box	23
8.4.2	Media Header Box.....	23
8.4.3	Handler Reference Box	24
8.4.4	Media Information Box	25
8.4.5	Media Information Header Boxes.....	25
8.5	Sample Tables.....	27
8.5.1	Sample Table Box.....	27
8.5.2	Sample Description Box	27
8.5.3	Degradation Priority Box	33
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.....	38
8.6.3	Shadow Sync Sample Box.....	39
8.6.4	Independent and Disposable Samples Box	40
8.6.5	Edit Box	41
8.6.6	Edit List Box	41
8.7	Track Data Layout Structures.....	43
8.7.1	Data Information Box	43
8.7.2	Data Reference Box.....	43
8.7.3	Sample Size Boxes	44
8.7.4	Sample To Chunk Box.....	45
8.7.5	Chunk Offset Box.....	46
8.7.6	Padding Bits Box	47
8.7.7	Sub-Sample Information Box	48
8.7.8	Sample Auxiliary Information Sizes Box	49
8.7.9	Sample Auxiliary Information Offsets Box	50
8.8	Movie Fragments	51
8.8.1	Movie Extends Box	51
8.8.2	Movie Extends Header Box	52
8.8.3	Track Extends Box	52
8.8.4	Movie Fragment Box	53
8.8.5	Movie Fragment Header Box	53
8.8.6	Track Fragment Box	54
8.8.7	Track Fragment Header Box	54
8.8.8	Track Fragment Run Box	55
8.8.9	Movie Fragment Random Access Box	56
8.8.10	Track Fragment Random Access Box	57
8.8.11	Movie Fragment Random Access Offset Box	58
8.8.12	Track fragment decode time	58
8.8.13	Level Assignment Box	59
8.8.14	Sample Auxiliary Information in Movie Fragments	61
8.9	Sample Group Structures	61
8.9.1	Introduction	61
8.9.2	Sample to Group Box	62
8.9.3	Sample Group Description Box	63
8.9.4	Representation of group structures in Movie Fragments	64
8.10	User Data	65
8.10.1	User Data Box	65
8.10.2	Copyright Box	65
8.10.3	Track Selection Box	66
8.11	Metadata Support.....	68
8.11.1	The Meta box	68
8.11.2	XML Boxes.....	69
8.11.3	The Item Location Box	69
8.11.4	Primary Item Box	71
8.11.5	Item Protection Box	72

8.11.6 Item Information Box	72
8.11.7 Additional Metadata Container Box	74
8.11.8 Metabox Relation Box	75
8.11.9 URL Forms for meta boxes	75
8.11.10 Static Metadata	76
8.11.11 Item Data Box	77
8.11.12 Item Reference Box	77
8.11.13 Auxiliary video metadata	78
8.12 Support for Protected Streams	78
8.12.1 Protection Scheme Information Box	79
8.12.2 Original Format Box	80
8.12.3 IPMPInfoBox	80
8.12.4 IPMP Control Box	80
8.12.5 Scheme Type Box	80
8.12.6 Scheme Information Box	81
8.13 File Delivery Format Support	81
8.13.1 Introduction	81
8.13.2 FD Item Information Box	82
8.13.3 File Partition Box	82
8.13.4 FEC Reservoir Box	84
8.13.5 FD Session Group Box	84
8.13.6 Group ID to Name Box	85
8.13.7 File Reservoir Box	86
8.14 Sub tracks	86
8.14.1 Introduction	86
8.14.2 Backward compatibility	87
8.14.3 Sub Track box	87
8.14.4 Sub Track Information box	87
8.14.5 Sub Track Definition box	88
8.14.6 Sub Track Sample Group box	89
8.15 Post-decoder requirements on media	89
8.15.1 General	89
8.15.2 Transformation	89
8.15.3 Restricted Scheme Information box	90
8.15.4 Scheme for stereoscopic video arrangements	90
8.16 Segments	92
8.16.1 Introduction	92
8.16.2 Segment Type Box	92
8.16.3 Segment Index Box	93
8.16.4 Subsegment Index Box	96
8.16.5 Producer Reference Time Box	97
9 Hint Track Formats	98
9.1 RTP and SRTP Hint Track Format	98
9.1.1 Introduction	98
9.1.2 Sample Description Format	99
9.1.3 Sample Format	101
9.1.4 SDP Information	104
9.1.5 Statistical Information	104
9.2 ALC/LCT and FLUTE Hint Track Format	105
9.2.1 Introduction	105
9.2.2 Design principles	106
9.2.3 Sample Description Format	107
9.2.4 Sample Format	108
9.3 MPEG-2 Transport Hint Track Format	111
9.3.1 Introduction	111
9.3.2 Design Principles	111
9.3.3 Sample Description Format	113
9.3.4 Sample Format	115
9.3.5 Protected MPEG 2 Transport Stream Hint Track	117

9.4	RTP, RTCP, SRTP and SRTCP Reception Hint Tracks	117
9.4.1	RTP Reception Hint Track.....	117
9.4.2	RTCP Reception Hint Track.....	121
9.4.3	SRTP Reception Hint Track	122
9.4.4	SRTCP Reception Hint Tracks.....	124
9.4.5	Protected RTP Reception Hint Track	125
9.4.6	Recording Procedure	125
9.4.7	Parsing Procedure	125
10	Sample Groups	125
10.1	Random Access Recovery Points.....	125
10.2	Rate Share Groups	126
10.2.1	Introduction	126
10.2.2	Rate Share Sample Group Entry	127
10.2.3	Relationship between tracks	128
10.2.4	Bitrate allocation.....	129
10.3	Alternative Startup Sequences.....	129
10.3.1	Definition	129
10.3.2	Syntax	130
10.3.3	Semantics	130
10.3.4	Examples	130
10.4	Random Access Point (RAP) Sample Grouping.....	132
10.4.1	Definition	132
10.4.2	Syntax	132
10.4.3	Semantics	132
10.5	Temporal level sample grouping.....	132
10.5.1	Definition	132
10.5.2	Syntax	133
10.5.3	Semantics	133
11	Extensibility.....	133
11.1	Objects.....	133
11.2	Storage formats	134
11.3	Derived File formats	134
Annex A (informative) Overview and Introduction		135
A.1	Section Overview.....	135
A.2	Core Concepts	135
A.3	Physical structure of the media	135
A.4	Temporal structure of the media	136
A.5	Interleave	136
A.6	Composition	136
A.7	Random access	137
A.8	Fragmented movie files.....	137
Annex B (informative) Patent Statements		139
Annex C (informative) Guidelines on deriving from this specification		140
C.1	Introduction	140
C.2	General Principles	140
C.2.1	General.....	140
C.2.2	Base layer operations	140
C.3	Boxes	141
C.4	Brand Identifiers	141
C.4.1	Introduction	141
C.4.2	Usage of the Brand	142
C.4.3	Introduction of a new brand	142
C.4.4	Player Guideline	142
C.4.5	Authoring Guideline	143
C.4.6	Example	143
C.5	Storage of new media types	143
C.6	Use of Template fields.....	144

C.7	Tracks	144
C.7.1	Data Location	144
C.7.2	Time	144
C.7.3	Media Types	145
C.7.4	Coding Types	145
C.7.5	Sub-sample information	145
C.7.6	Sample Dependency	145
C.7.7	Sample Groups	145
C.7.8	Track-level	145
C.7.9	Protection	146
C.8	Construction of fragmented movies	146
C.9	Meta-data	147
C.10	Registration	147
C.11	Guidelines on the use of sample groups, timed metadata tracks, and sample auxiliary information	147
Annex D (informative)	Registration Authority	149
D.1	Code points to be registered	149
D.2	Procedure for the request of an MPEG-4 registered identifier value	149
D.3	Responsibilities of the Registration Authority	150
D.4	Contact information for the Registration Authority	150
D.5	Responsibilities of Parties Requesting a RID	150
D.6	Appeal Procedure for Denied Applications	151
D.7	Registration Application Form	151
D.7.1	Contact Information of organization requesting a RID	151
D.7.2	Request for a specific RID	151
D.7.3	Short description of RID that is in use and date system was implemented	152
D.7.4	Statement of an intention to apply the assigned RID	152
D.7.5	Date of intended implementation of the RID	152
D.7.6	Authorized representative	152
D.7.7	For official use of the Registration Authority	152
Annex E (normative)	File format brands	153
E.1	Introduction	153
E.2	The 'isom' brand	154
E.3	The 'avc1' brand	155
E.4	The 'iso2' brand	155
E.5	The 'mp71' brand	156
E.6	The 'iso3' brand	156
E.7	The 'iso4' brand	156
E.8	The 'iso5' brand	157
E.9	The 'iso6' brand	157
Annex F (informative)	Document Cross-Reference	158
Annex G (informative)	URI-labelled metadata forms	160
G.1	UUID-labelled metadata	160
G.2	ISO OID-labelled metadata	160
G.3	SMPTE-labelled metadata	160
Annex H (informative)	Processing of RTP streams and reception hint tracks	162
H.1	Introduction	162
H.1.1	Overview	162
H.1.2	Structure	162
H.1.3	Terms and definitions	162
H.2	Synchronization of RTP streams	162
H.3	Recording of RTP streams	163
H.3.1	Introduction	163
H.3.2	Compensation for unequal starting for position of received RTP streams	165
H.3.3	Recording of SDP	166
H.3.4	Creation of a sample within an RTP reception hint track	166
H.3.5	Representation of RTP timestamps	167

H.3.6 Recording operations to facilitate inter-stream synchronization in playback	170
H.3.7 Representation of reception times.....	171
H.3.8 Creation of media samples	172
H.3.9 Creation of hint samples referring to media samples.....	172
H.4 Playing of recorded RTP streams	172
H.4.1 Introduction	172
H.4.2 Preparation for the playback	173
H.4.3 Decoding of a sample within an RTP reception hint track	173
H.4.4 Lip synchronization.....	173
H.4.5 Random access.....	175
H.5 Re-sending recorded RTP streams.....	175
H.5.1 Introduction	175
H.5.2 Re-sending RTP packets.....	176
H.5.3 RTCP Processing.....	177
Annex I (normative) Stream Access Points.....	178
I.1 Introduction	178
I.2 SAP properties	178
I.3 SAP types	178
Annex J (normative) MIME Type Registration of Segments	180
J.1 Introduction	180
J.2 Registration	180
Bibliography	181



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

A large, stylized, red-outlined watermark reading "WATERMARK" diagonally across the page.

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 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.)*