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

**Part 15:
Carriage of network abstraction layer
(NAL) unit structured video in the ISO
base media file format**

Technologies de l'information — Codage des objets audiovisuels —

*Partie 15: Transport de vidéo structuré en unités NAL au format ISO
de base pour les fichiers médias*

This document is a preview generated by EBS



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword	vi
Introduction	vii
1 Scope	1
2 Normative references	1
3 Terms, definitions and abbreviated terms	1
3.1 Terms and definitions	1
3.2 Abbreviated terms	7
4 General definitions	8
4.1 Overview	8
4.2 Elementary stream structure	8
4.3 Sample and configuration definition	8
4.3.1 General	8
4.3.2 Canonical order and restrictions	9
4.3.3 Sample format	9
4.3.4 Optional boxes in the sample entry	10
4.4 Video track structure	10
4.5 Template fields used	11
4.6 Visual width and height	11
4.7 Decoding time (DTS) and composition time (CTS)	11
4.8 Sample groups on random access recovery points 'roll' and random access points 'rap'	11
4.9 Hinting	12
4.10 On change of sample entry	12
5 AVC elementary streams and sample definitions	13
5.1 General	13
5.2 Elementary stream structure	14
5.3 Sample and configuration definition	16
5.3.1 General	16
5.3.2 Canonical order and restrictions	16
5.3.3 Decoder configuration information	17
5.4 Derivation from ISO base media file format	20
5.4.1 AVC file type and identification	20
5.4.2 AVC video stream definition	20
5.4.3 AVC parameter set stream definition	21
5.4.4 Parameter sets	22
5.4.5 Sync sample	23
5.4.6 Shadow sync	23
5.4.7 Layering and sub-sequences	23
5.4.8 Alternate streams and switching pictures	27
5.4.9 Definition of a sub-sample for AVC	29
5.4.10 Post-decoder requirements scheme for signalling of SEI for AVC	29
6 SVC elementary stream and sample definitions	30
6.1 General	30
6.2 Elementary stream structure	31
6.3 Use of the plain AVC file format	32
6.4 Sample and configuration definition	32
6.4.1 General	32
6.4.2 Canonical order and restrictions	32
6.5 Derivation from the ISO base media file format	34
6.5.1 SVC track structure	34
6.5.2 Data sharing and extraction	34
6.5.3 SVC video stream definition	35

6.5.4	SVC visual width and height	37
6.5.5	Sync sample	37
6.5.6	Shadow sync	37
6.5.7	Independent and disposable samples box	37
6.5.8	Sample groups on random access recovery points 'roll' and random access points 'rap'	38
6.5.9	Definition of a sub-sample for SVC	38
7	MVC and MVD elementary stream and sample definitions	39
7.1	General	39
7.2	Overview of MVC or MVD Storage	40
7.3	MVC and MVD elementary stream structures	42
7.4	Use of the plain AVC file format	43
7.5	Sample and configuration definition	44
7.5.1	General	44
7.5.2	Canonical order and restriction	44
7.5.3	Decoder configuration record	44
7.6	Derivation from the ISO base media file format	46
7.6.1	MVC and MVD track structures	46
7.6.2	Reconstruction of an access unit	47
7.6.3	Sample entry	48
7.6.4	Sync sample	58
7.6.5	Shadow sync	59
7.6.6	Independent and disposable samples box	59
7.6.7	Sample groups on random access recovery points 'roll' and random access points 'rap'	59
7.7	MVC specific information boxes	59
7.7.1	General	59
7.7.2	Multiview information box	60
7.7.3	Multiview group box	60
7.7.4	Multiview group relation box	62
7.7.5	Multiview relation attribute box	62
7.7.6	Multiview scene info box	67
7.7.7	MVC view priority assignment box	68
8	HEVC elementary streams and sample definitions	68
8.1	General	68
8.2	Elementary stream structure	69
8.3	Sample and configuration definition	69
8.3.1	General	69
8.3.2	Canonical order and restrictions	69
8.3.3	Decoder configuration information	70
8.4	Derivation from ISO base media file format	73
8.4.1	HEVC video stream definition	73
8.4.2	Parameter sets in sample entry	74
8.4.3	Sync sample	74
8.4.4	Sync sample sample grouping	75
8.4.5	Temporal scalability sample grouping	75
8.4.6	Temporal sub-layer access sample grouping	77
8.4.7	Step-wise temporal layer access sample grouping	77
8.4.8	Definition of a sub-sample for HEVC	78
8.4.9	Handling non-output samples	80
9	Layered HEVC elementary stream and sample definitions	80
9.1	General	80
9.2	Overview of L-HEVC storage	81
9.3	L-HEVC elementary stream structure	82
9.4	Sample and configuration definition	82
9.4.1	General	82
9.4.2	Canonical order and restrictions	82

9.4.3	Decoder configuration record	82
9.5	Derivation from the ISO base media file format and the HEVC file format (Clause 8)	83
9.5.1	L-HEVC track structure	83
9.5.2	Data sharing and reconstruction of an L-HEVC bitstream	84
9.5.3	L-HEVC video stream definition	85
9.5.4	L-HEVC visual width and height	88
9.5.5	Sync sample	88
9.5.6	Independent and disposable samples box	89
9.5.7	Stream access point sample group	89
9.5.8	The 'roll', 'rap', 'sync', 'tsas' and 'stsa' sample groups	89
9.5.9	Definition of a sub-sample for L-HEVC	90
9.5.10	Handling non-output samples	90
9.6	L-HEVC specific structures	90
9.6.1	External base layer sample group	90
9.6.2	The operating points information sample group	91
9.6.3	The layer information sample group	94
10	Storage of tiled HEVC and L-HEVC video streams	95
10.1	General	95
10.2	NAL unit map entry	95
10.2.1	Definition	95
10.2.2	Syntax	96
10.2.3	Semantics	96
10.3	Tile region group entry	97
10.3.1	Definition	97
10.3.2	Syntax	97
10.3.3	Semantics	97
10.4	Tile sub track definition	99
10.4.1	General	99
10.4.2	TileSubTrackGroupBox	99
10.5	HEVC and L-HEVC tile track	99
10.5.1	General	99
10.5.2	Sample entry name and format for HEVC tile tracks	100
10.5.3	Sample entry name and format for L-HEVC tile tracks	101
10.5.4	Bitstream reconstruction from tile base and tile tracks	102
10.5.5	Sample entry names for tile base tracks	102
	Annex A (normative) In-stream structures	104
	Annex B (normative) SVC, MVC, and MVD sample group and sub-track definitions	113
	Annex C (normative) Temporal metadata support	134
	Annex D (normative) File format toolsets and brands	143
	Annex E (normative) Sub-parameters for the MIME type "codecs" parameter	145

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.

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 ISO documents 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 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).

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 voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is 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 14496-15:2014), which has been technically revised.

It also incorporates the Technical Corrigendum ISO/IEC 14496-15:2014/Cor 1:2015.

A list of all parts in the ISO/IEC 14496 series can be found on the ISO website.

Introduction

This document defines a storage format based on, and compatible with, the ISO Base Media File Format (ISO/IEC 14496-12), which is used by the MP4 file format (ISO/IEC 14496-14) and the Motion JPEG 2000 file format (ISO/IEC 15444-3) among others. This document enables video streams formatted as Network Adaptation Layer Units (NAL Units) to

- a) be used in conjunction with other media streams, such as audio,
- b) be used in an MPEG-4 systems environment, if desired,
- c) be formatted for delivery by a streaming server, using hint tracks, and
- d) inherit all the use cases and features of the ISO Base Media File Format on which MP4 and MJ2 are based.

This document may be used as a standalone specification; it specifies how NAL unit structured video content shall be stored in an ISO Base Media File Format compliant format. However, it is normally used in the context of a specification, such as the MP4 file format, derived from the ISO Base Media File Format, that permits the use of NAL unit structured video such as AVC (ISO/IEC 14496-10) video and High Efficiency Video Coding (HEVC, ISO/IEC 23008-2) video.

The ISO Base Media File Format is becoming increasingly common as a general-purpose media container format for the exchange of digital media, and its use in this context should accelerate both adoption and interoperability.

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 a patent.

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.

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.

Information technology — Coding of audio-visual objects —

Part 15:

Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format

1 Scope

This document specifies the storage format for streams of video that is structured as NAL units, such as AVC (ISO/IEC 14496-10) and HEVC (ISO/IEC 23008-2) video streams.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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/IEC 14496-12:2015, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 14496-10:2014, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

ISO/IEC 23008-2:—¹⁾, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-10 or ISO/IEC 23008-2, and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at <http://www.electropedia.org/>
- ISO Online browsing platform: available at <http://www.iso.org/obp>

3.1.1

3D-AVC NAL unit

3D-AVC VCL NAL unit

NAL unit with type 21 with `avc_3d_extension_flag` equal to 1

Note 1 to entry: As specified in ISO/IEC 14496-10:2014, Annex J.

1) To be published.