ISO/IEC 14496-15:2019 (E)

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

Contents

Foreword

Introduction

1 Scope

2 Normative references

3 Terms, definitions and abbreviated terms

3.1 Terms and definitions
3.2 Abbreviated terms

4 General definitions

4.1 Overview
4.2 Elementary stream structure
4.3 Sample and configuration definition
4.3.1 General
4.3.2 Canonical order and restrictions
4.3.3 Sample format
4.3.3.1 Definition
4.3.3.2 Syntax
4.3.3.3 Semantics
4.3.4 Optional boxes in the sample entry
4.4 Video track structure
4.5 Template fields used
4.6 Visual width and height
4.7 Decoding time (DTS) and composition time (CTS)
4.8 Sample groups on random access recovery points 'roll' and random access points 'rap '
4.9 Hinting
4.10 On change of sample entry
4.11 SEI information box
4.11.1 Definition
4.11.2 Syntax
4.11.3 Semantics
4.12 Post-decoder requirements scheme for signalling of SEI
4.12.1 General
4.12.2 Definition

5 AVC elementary streams and sample definitions

5.1 Overview
5.2 Elementary stream structure
5.3 Sample and configuration definition
5.3.1 Overview
5.3.2 Canonical order and restrictions
5.3.3 Decoder configuration information
5.3.3.1 AVC decoder configuration record
5.3.3.1.1 Definition
5.3.3.1.2 Syntax
5.3.3.1.3 Semantics
5.4 Derivation from ISO base media file format
5.4.1 AVC file type and identification

5.4.2 AVC video stream definition
5.4.2.1 Sample entry name and format
5.4.2.1.1 Definition
5.4.2.1.2 Syntax
5.4.2.1.3 Semantics
5.4.3 AVC parameter set stream definition
5.4.3.1 Sample entry name and format
5.4.3.1.1 Definition
5.4.3.1.2 Syntax
5.4.3.1.3 Semantics
5.4.3.2 Sample format
5.4.3.2.1 Definition
5.4.3.2.2 Syntax
5.4.3.2.3 Semantics
5.4.3.3 Track reference
5.4.4 Parameter sets
5.4.5 Sync sample
5.4.6 Shadow sync
5.4.7 Layering and sub-sequences
5.4.7.1 Overview
5.4.7.2 Sub-sequence description entry
5.4.7.2.1 Definition
5.4.7.2.2 Syntax
5.4.7.2.3 Semantics
5.4.7.3 Layer description entry
5.4.7.3.1 Definition
5.4.7.3.2 Syntax
5.4.7.3.3 Semantics
5.4.8 Alternate streams and switching pictures
5.4.8.1 Switching pictures
5.4.8.2 Alternate group
5.4.8.3 Track references
5.4.8.4 Sample dependency
5.4.8.4.1 Definition
5.4.8.4.2 Syntax
5.4.8.4.3 Semantics
5.4.9 Definition of a sub-sample for AVC

6 SVC elementary stream and sample definitions

6.1 Overview
6.2 Elementary stream structure
6.3 Use of the plain AVC file format
6.4 Sample and configuration definition
6.4.1 Overview
6.4.2 Canonical order and restrictions
6.4.2.1 Restrictions
6.4.2.2 Decoder configuration record
6.5 Derivation from the ISO base media file format
6.5.1 SVC track structure
6.5.2 Data sharing and extraction
6.5.3 SVC video stream definition
6.5.3.1 Sample entry name and format
6.5.3.1.1 Definition
6.5.3.1.2 Syntax
6.5.3.1.3 Semantics
6.5.4 SVC visual width and height
6.5.5 Sync sample
6.5.6 Shadow sync
6.5.7 Independent and disposable samples box
6.5.8 Sample groups on random access recovery points 'roll' and random access points 'rap '

6.5.9 Definition of a sub-sample for SVC

7 MVC and MVD elementary stream and sample definitions

7.1 Overview
7.2 Overview of MVC or MVD storage
7.3 MVC and MVD elementary stream structures
7.4 Use of the plain AVC file format
7.5 Sample and configuration definition
7.5.1 Overview
7.5.2 Canonical order and restriction
7.5.3 Decoder configuration record
7.5.3.1 MVC decoder configuration record
7.5.3.2 MVD decoder configuration record
7.6 Derivation from the ISO base media file format
7.6.1 MVC and MVD track structures
7.6.2 Reconstruction of an access unit
7.6.3 Sample entry
7.6.3.1 Boxes for sample entry
7.6.3.1.1 Intrinsic camera parameters box
7.6.3.1.1.1 Definition
7.6.3.1.1.2 Syntax
7.6.3.1.1.3 Semantics
7.6.3.1.2 Extrinsic camera parameters box
7.6.3.1.2.1 Definition
7.6.3.1.2.2 Syntax
7.6.3.1.2.3 Semantics
7.6.3.1.3 View identifier box
7.6.3.1.3.1 Definition
7.6.3.1.3.2 Syntax
7.6.3.1.3.3 Semantics
7.6.3.2 MVC and MVD sample entry definitions
7.6.3.2.1 Definition
7.6.3.2.2 Syntax
7.6.3.2.3 Semantics
7.6.4 Sync sample
7.6.5 Shadow sync
7.6.6 Independent and disposable samples box
7.6.7 Sample groups on random access recovery points 'roll' and random access points 'rap '

7.7 MVC specific information boxes
7.7.1 Overview
7.7.2 Multiview information box
7.7.2.1 Definition
7.7.2.2 Syntax
7.7.3 Multiview group box
7.7.3.1 Definition
7.7.3.2 Syntax
7.7.3.3 Semantics
7.7.4 Multiview group relation box
7.7.4.1 Definition
7.7.4.2 Syntax
7.7.4.3 Semantics
7.7.5 Multiview relation attribute box
7.7.5.1 Definition
7.7.5.2 Syntax
7.7.5.3 Semantics
7.7.6 Multiview scene info box
7.7.6.1 Definition
7.7.6.2 Syntax
7.7.6.3 Semantics
7.7.7 MVC view priority assignment box
7.7.7.1 Definition
7.7.7.2 Syntax
7.7.7.3 Semantics

8 HEVC elementary streams and sample definitions

8.1 Overview

8.2 Elementary stream structure
8.3 Sample and configuration definition
8.3.1 Overview
8.3.2 Canonical order and restrictions
8.3.3 Decoder configuration information
8.3.3.1 HEVC decoder configuration record
8.3.3.1.1 Definition
8.3.3.1.2 Syntax
8.3.3.1.3 Semantics
8.4 Derivation from ISO base media file format
8.4.1 HEVC video stream definition
8.4.1.1 Sample entry name and format
8.4.1.1.1 Definition
8.4.1.1.2 Syntax
8.4.1.1.3 Semantics
8.4.2 Parameter sets in sample entry
8.4.3 Sync sample
8.4.4 Sync sample sample grouping
8.4.4.1 Overview
8.4.4.2 Sync sample sample group entry
8.4.4.2.1 Definition
8.4.4.2.2 Syntax
8.4.4.2.3 Semantics
8.4.5 Temporal scalability sample grouping
8.4.5.1 Overview
8.4.5.2 Temporal layer sample group entry
8.4.5.2.1 Definition
8.4.5.2.2 Syntax
8.4.5.2.3 Semantics
8.4.6 Temporal sub-layer access sample grouping
8.4.6.1 Overview
8.4.6.2 Temporal sub-layer access sample group entry
8.4.6.2.1 Overview
8.4.6.2.2 Syntax
8.4.7 Step-wise temporal layer access sample grouping
8.4.7.1 Overview
8.4.7.2 Step-wise temporal layer sample group entry
8.4.7.2.1 Definition
8.4.7.2.2 Syntax
8.4.8 Definition of a sub-sample for HEVC
8.4.9 Handling non-output samples

9 Layered HEVC elementary stream and sample definitions

9.1 Overview
9.2 Overview of L-HEVC storage
9.3 L-HEVC elementary stream structure
9.4 Sample and configuration definition
9.4.1 Overview
9.4.2 Canonical order and restrictions
9.4.3 Decoder configuration record
9.5 Derivation from the ISO base media file format and the HEVC file format (Clause 8)
9.5.1 L-HEVC track structure
9.5.2 Data sharing and reconstruction of an L-HEVC bitstream
9.5.2.1 General
9.5.2.2 Implicit reconstruction
9.5.3 L-HEVC video stream definition
9.5.3.1 Sample entry name and format
9.5.3.1.1 Definition
9.5.3.1.2 Syntax
9.5.3.1.3 Semantics
9.5.4 L-HEVC visual width and height
9.5.5 Sync sample
9.5.6 Independent and disposable samples box
9.5.7 Stream access point sample group

9.5.8 The 'roll', 'rap ', 'sync', 'tsas' and 'stsa' sample groups
9.5.9 Definition of a sub-sample for L-HEVC
9.5.10 Handling non-output samples
9.6 L-HEVC specific structures
9.6.1 External base layer sample group
9.6.1.1 Definition
9.6.1.2 Syntax
9.6.1.3 Semantics
9.6.2 The operating points information sample group
9.6.2.1 Definition
9.6.2.2 Syntax
9.6.2.3 Semantics
9.6.3 The layer information sample group
9.6.3.1 Definition
9.6.3.2 Syntax
9.6.3.3 Semantics
9.6.4 The layer information sample group
9.6.4.1 Definition
9.6.4.2 Syntax
9.6.4.3 Semantics

10 Storage of tiled HEVC and L-HEVC video streams

10.1 Overview
10.2 NAL unit map entry
10.2.1 Definition
10.2.2 Syntax
10.2.3 Semantics
10.3 Tile region group entry
10.3.1 Definition
10.3.2 Syntax
10.3.3 Semantics
10.4 Tile sub track definition
10.4.1 Overview
10.4.2 TileSubTrackGroupBox
10.4.2.1 Definition
10.4.2.2 Syntax
10.4.2.3 Semantics
10.5 HEVC and L-HEVC tile track
10.5.1 Overview
10.5.2 Sample entry name and format for HEVC tile tracks
10.5.2.1 Definition
10.5.2.2 Syntax
10.5.2.3 Semantics
10.5.3 Sample entry name and format for L-HEVC tile tracks
10.5.3.1 Definition
10.5.3.2 Syntax
10.5.3.3 Semantics
10.5.4 Bitstream reconstruction from tile base and tile tracks
10.5.5 Sample entry names for tile base tracks

Annex A (normative) In-stream structures

A.1 General
A.2 Aggregators
A.2.1 Definition
A.2.2 Syntax
A.2.3 Semantics
A.3 Extractors for SVC, MVC, and MVD tracks
A.3.1 Definition
A.3.2 Syntax
A.3.3 Semantics
A.4 NAL unit header values for SVC
A.5 NAL unit header values for MVC and MVC+D depth NAL units
A.6 NAL unit header values for 3D-AVC NAL units
A.7 Extractors for HEVC and L-HEVC tracks

A.7.1 Definition
A.7.2 Syntax
A.7.3 Semantics
A.7.4 Sample constructor
A.7.4.1 Syntax
A.7.4.2 Semantics
A.7.5 In-line constructor
A.7.5.1 Syntax
A.7.5.2 Semantics
A.8 NAL unit header values for ISO/IEC 23008-2

Annex B (normative) SVC, MVC, and MVD sample group and sub-track definitions

B.1 General
B.2 Definition
B.2.1 Tier information box
B.2.1.1 Definition
B.2.1.2 Syntax
B.2.1.3 Semantics
B.2.2 Tier bit rate box
B.2.2.1 Definition
B.2.2.2 Syntax
B.2.2.3 Semantics
B.2.3 Priority range
B.2.3.1 Definition
B.2.3.2 Syntax
B.2.3.3 Semantics
B.2.4 SVC dependency range
B.2.4.1 Definition
B.2.4.2 Syntax
B.2.4.3 Semantics
B.2.5 Initial parameter sets box
B.2.5.1 Definition
B.2.5.2 Syntax
B.2.5.3 Semantics
B.2.6 SVC rect region box
B.2.6.1 Definition
B.2.6.2 Syntax
B.2.6.3 Semantics
B.2.7 Buffering information box
B.2.7.1 Definition
B.2.7.2 Syntax
B.2.7.3 Semantics
B.2.8 Tier dependency box
B.2.8.1 Definition
B.2.8.2 Syntax
B.2.8.3 Semantics
B.2.9 SVC region of interest box
B.2.9.1 Definition
B.2.9.2 Syntax
B.2.9.3 Semantics
B.2.10 SVC lightweight transcoding Box
B.2.10.1 Definition
B.2.10.2 Syntax
B.2.10.3 Semantics
B.2.11 Scalable and multiview group entries
B.2.11.1 Overview
B.2.11.2 Scalable group entry
B.2.11.2.1 Definition
B.2.11.2.2 Syntax
B.2.11.2.3 Semantics
B.2.11.3 Multiview group entry
B.2.11.3.1 Definition
B.2.11.3.2 Syntax
B.2.11.3.3 Semantics

B.3 Mapping NAL units to map groups and tiers
B.3.1 Overview
B.3.2 Map group definition
B.4 Decode re-timing groups
B.4.1 Overview
B.4.2 Syntax
B.4.3 Semantics
B.5 View priority sample grouping
B.5.1 Definition
B.5.2 Syntax
B.5.3 Semantics
B.6 Sub track definitions
B.6.1 General
B.6.2 SVC sub track layer box
B.6.2.1 Definition
B.6.2.2 Syntax
B.6.2.3 Semantics
B.6.3 MVC sub track view box
B.6.3.1 Definition
B.6.3.2 Syntax
B.6.3.3 Semantics
B.6.4 Sub track tier box
B.6.4.1 Definition
B.6.4.2 Syntax
B.6.4.3 Semantics
B.6.5 MVC sub track multiview group box
B.6.5.1 Definition
B.6.5.2 Syntax
B.6.5.3 Semantics

Annex C (normative) Temporal metadata support

C.1 General
C.2 Connection to the video media data
C.3 SVC meta data sample entry
C.3.1 Definition
C.3.2 Syntax
C.3.3 Semantics
C.4 Helper functions

Annex D (normative) File format toolsets and brands

D.1 General
D.2 SVC Toolsets
D.3 MVC and MVD toolsets
D.4 L-HEVC brands
D.4.1 L-HEVC explicit reconstruction brand
D.4.2 L-HEVC implicit reconstruction brand
D.4.3 HEVC Tile Track brand
D.4.4 L-HEVC Tile Track Implicit brand
D.4.5 L-HEVC Tile Track Explicit brand
D.5 No Leading Picture Sync Brand

Annex E (normative) Sub-parameters for the MIME type ‘codecs’ parameter

E.1 General
E.2 AVC family
E.3 HEVC
E.4 L-HEVC
E.5 HEVC and L-HEVC tile tracks

Annex F (informative) Unspecified nal_unit_type value management

Page count: 171