ISO/IEC 14496-15:2024-10 (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

Page

Foreword ... **viii**

Introduction ... **ix**

1 Scope ... **1**

2 Normative references ... **1**

3 Terms, definitions, abbreviated terms and conventions ... **1**
3.1 Terms and definitions ... 1
3.2 Abbreviated terms ... 9
3.3 Conventions ... 11

4 General definitions ... **11**
4.1 Overview ... 11
4.2 Sample and configuration definition ... 11
4.2.1 General ... 11
4.2.2 Canonical order and restrictions ... 11
4.2.3 Sample format ... 12
4.2.4 Optional boxes in the sample entry ... 13
4.3 Video track structure ... 13
4.4 Template fields used ... 13
4.5 Visual width and height ... 13
4.6 Decoding time (DTS) and composition time (CTS) ... 14
4.7 Sample groups on random access recovery points `'roll'` and random access points `'rap '` ... 14
4.8 Hinting ... 14
4.9 On change of sample entry (informative) ... 15
4.10 SEI information box ... 16
4.10.1 Definition ... 16
4.10.2 Syntax ... 16
4.10.3 Semantics ... 16
4.11 Post-decoder requirements scheme for signalling of SEI ... 16
4.11.1 General ... 16
4.11.2 Definition ... 17
4.12 Alternative extraction source track grouping ... 17
4.13 NAL unit map entry ... 17
4.13.1 Definition ... 17
4.13.2 Syntax ... 18
4.13.3 Semantics ... 18
4.14 Rectangular region group entry ... 19
4.14.1 Definition ... 19
4.14.2 Syntax ... 19
4.14.3 Semantics ... 19
4.15 Layer information sample group ... 21
4.15.1 Definition ... 21
4.15.2 Syntax ... 21
4.15.3 Semantics ... 21
4.16 Storage of SEI manifest and SEI prefix indication SEI messages ... 22
4.17 Supplementary track reference ... 22
4.18 Picture region replacement sample group ... 23
4.18.1 Definition ... 23

4.18.2 Syntax 23
4.18.3 Semantics 23

5 AVC elementary streams and sample definitions 24
5.1 Overview 24
5.2 Elementary stream structure 25
5.3 Sample and configuration definition 27
5.3.1 Canonical order and restrictions 27
5.3.2 Decoder configuration information 28
5.4 Derivation from ISO base media file format 30
5.4.1 AVC file type and identification 30
5.4.2 AVC video stream definition 31
5.4.3 AVC parameter set stream definition 32
5.4.4 Parameter sets 33
5.4.5 Sync sample 34
5.4.6 Shadow sync 34
5.4.7 Layering and sub-sequences 34
5.4.8 Alternate streams and switching pictures 38
5.4.9 Definition of a sub-sample for AVC 40

6 SVC elementary stream and sample definitions 40
6.1 Overview 40
6.2 Elementary stream structure 41
6.3 Use of the plain AVC file format 41
6.4 Sample and configuration definition 42
6.4.1 Canonical order and restrictions 42
6.4.2 Decoder configuration record 42
6.5 Derivation from the ISO base media file format 43
6.5.1 SVC track structure 43
6.5.2 Data sharing and extraction 43
6.5.3 SVC video stream definition 44
6.5.4 SVC visual width and height 46
6.5.5 Sync sample 46
6.5.6 Shadow sync 46
6.5.7 Independent and disposable samples box 47
6.5.8 Sample groups on random access recovery points `'roll'` and random access points `'rap '` 47
6.5.9 Definition of a sub-sample for SVC 47

7 MVC and MVD elementary stream and sample definitions 48
7.1 Overview 48
7.2 Overview of MVC or MVD Storage 49
7.3 MVC and MVD elementary stream structures 51
7.4 Use of the plain AVC file format 52
7.5 Sample and configuration definition 52
7.5.1 Canonical order and restriction 52
7.5.2 Decoder configuration record 53
7.6 Derivation from the ISO base media file format 55
7.6.1 MVC and MVD track structures 55
7.6.2 Reconstruction of an access unit 55
7.6.3 Sample entry 56
7.6.4 Sync sample 66
7.6.5 Shadow sync 66
7.6.6 Independent and disposable samples box 66
7.6.7 Sample groups on random access recovery points `'roll'` and random access points `'rap '` 66
7.7 MVC specific information boxes 67
7.7.1 Overview 67
7.7.2 Multiview information box 67
7.7.3 Multiview group box 67
7.7.4 Multiview group relation box 69
7.7.5 Multiview relation attribute box 70
7.7.6 Multiview scene info box 74
7.7.7 MVC view priority assignment box 75

8 HEVC elementary streams and sample definitions 75
8.1 Overview 75

8.2 Elementary stream structure 76
8.3 Sample and configuration definition 76
8.3.1 Canonical order and restrictions 76
8.3.2 Decoder configuration information 77
8.4 Derivation from ISO base media file format 80
8.4.1 HEVC video stream definition 80
8.4.2 Parameter sets in sample entry 81
8.4.3 Sync sample 81
8.4.4 Sync sample sample grouping 82
8.4.5 Temporal scalability sample grouping 83
8.4.6 Temporal sub-layer access sample grouping 84
8.4.7 Step-wise temporal layer access sample grouping 85
8.4.8 Definition of a sub-sample for HEVC 85
8.4.9 Handling non-output samples 88

9 Layered HEVC elementary stream and sample definitions 88
9.1 Overview 88
9.2 Overview of L-HEVC storage 89
9.3 L-HEVC elementary stream structure 89
9.4 Sample and configuration definition 90
9.4.1 Overview 90
9.4.2 Canonical order and restrictions 90
9.4.3 Decoder configuration record 90
9.5 Derivation from the ISO base media file format and the HEVC file format (Clause 8) 91
9.5.1 L-HEVC track structure 91
9.5.2 Data sharing and reconstruction of an L-HEVC bitstream 92
9.5.3 L-HEVC video stream definition 93
9.5.4 L-HEVC visual width and height 97
9.5.5 Sync sample 97
9.5.6 Independent and disposable samples box 98
9.5.7 Stream access point sample group 98
9.5.8 The `'roll'`, `'rap '`, `'sync'`, `'tsas'` and `'stsa'` sample groups 98
9.5.9 Definition of a sub-sample for L-HEVC 99
9.5.10 Handling non-output samples 99
9.6 L-HEVC specific structures 99
9.6.1 External base layer sample group 99
9.6.2 The operating points information sample group 100
9.6.3 The layer information sample group 103
9.6.4 The decoding time hint sample group 103

10 Storage of tiled HEVC and L-HEVC video streams 104
10.1 Overview 104
10.2 NAL unit map entry 105
10.3 Tile region group entry 105
10.4 Tile sub track definition 105
10.4.1 Overview 105
10.4.2 TileSubTrackGroupBox 105
10.5 HEVC and L-HEVC tile track 106
10.5.1 Overview 106
10.5.2 Sample entry name and format for HEVC tile tracks 107
10.5.3 Sample entry name and format for L-HEVC tile tracks 108
10.5.4 Bitstream reconstruction from tile base and tile tracks 108
10.5.5 Sample entry names for tile base tracks 109
10.5.6 HEVC tile track with slice segment header info 109
10.6 HEVC slice segment data track 110
10.6.1 Overview 110
10.6.2 Sample entry name and format for HEVC slice segment data tracks 110

11 VVC elementary streams and sample definitions 111
11.1 Overview 111

11.1.1 General 111
11.1.2 Background: features of VVC 111
11.1.3 Types of tracks for carriage of VVC elementary streams 112
11.1.4 Overview of VVC storage with multiple layers or sublayers 113
11.1.5 Overview of VVC storage with VVC subpictures 114
11.1.6 Overview of rectangular regions carried in a VVC bitstream 115
11.2 Sample and configuration definition 116
11.2.1 Sample format of VVC tracks and VVC subpicture tracks 116
11.2.2 Sample format of VVC non-VCL tracks 116
11.2.3 Canonical order and restrictions 117
11.2.4 Decoder configuration information 118
11.3 Derivation from ISO base media file format 124
11.3.1 VVC sample entries 124
11.3.2 VVC subpicture sample entry `'vvs1'` 125
11.3.3 VVC non-VCL sample entry 126
11.3.4 Constraints related to VVC merge base tracks, VVC extraction base tracks and VVC subpicture tracks 127
11.3.5 Sync sample 128
11.3.6 Definition of a sub-sample for VVC 131
11.3.7 Handling non-output samples 134
11.4 Sample groups 134
11.4.1 Common layer_id_method_idc semantics 134
11.4.2 Stream access point sample group 136
11.4.3 Random access recovery point sample group 136
11.4.4 Alternative startup sequences sample group 136
11.4.5 Random access point sample group 136
11.4.6 Temporal level sample group 136
11.4.7 Step-wise sublayer access sample group 137
11.4.8 Decoding time hint sample group 137
11.4.9 Layer information sample group 137
11.4.10 Operating points information sample group 137
11.4.11 Decoding capability information sample group 142
11.4.12 Parameter set sample group 142
11.4.13 Access unit delimiter sample group 143
11.4.14 End of sequence sample group 144
11.4.15 End of bitstream sample group 144
11.4.16 Subpicture ID sample group 145
11.4.17 Subpicture order sample group 146
11.4.18 Subpicture layout map entry 147
11.4.19 Mixed NAL unit type pictures sample group 148
11.4.20 Rectangular region order sample group 149
11.4.21 Subpicture level information sample group 150
11.5 Entity groups 151
11.5.1 Subpicture entity groups 151
11.5.2 Operating point entity group 153
11.5.3 VVC bitstream entity group 156
11.5.4 VVC switchable tracks entity group 156
11.6 Data sharing and VVC bitstream reconstruction 157
11.6.1 General 157
11.6.2 Implicit reconstruction of a VVC bitstream 159
11.6.3 Reconstructing a picture unit from a sample in a VVC track with `'subp'` or `'vvcN'` track references 160
11.6.4 Resolving subpicture track references 162
11.6.5 Parameter set updating 162
11.6.6 Reconstructing a picture unit from a sample in a VVC track with `'recr'` track reference 163

12 EVC elementary streams and sample definitions 165
12.1 Overview 165

12.2 Elementary stream structure ... 165
12.3 Sample and configuration definition ... 165
12.3.1 Overview ... 165
12.3.2 Canonical order and restrictions ... 165
12.3.3 Decoder configuration information: EVC decoder configuration record ... 166
12.4 Derivation from ISO base media file format ... 168
12.4.1 EVC video stream definition: sample entry name and format ... 168
12.4.2 Parameter sets ... 169
12.4.3 Sync sample ... 170
12.4.4 Definition of a sub-sample for EVC ... 170
12.5 EVC slice track ... 171
12.5.1 Overview ... 171
12.5.2 Implicit reconstruction of an EVC bitstream ... 171
12.5.3 EVC slice component track ... 171
12.5.4 EVC slice base track ... 173

13 LCEVC elementary streams and sample definitions ... 173
13.1 Overview ... 173
13.2 Elementary stream structure ... 174
13.3 Sample and configuration definitions ... 174
13.3.1 Overview ... 174
13.3.2 Canonical order ... 174
13.3.3 Decoder configuration information ... 174
13.4 Derivation from ISO base media file format ... 176
13.4.1 LCEVC video stream definition: sample entry name and format ... 176
13.4.2 LCEVC track structure ... 177
13.4.3 Parameter sets ... 178
13.4.4 Sync sample ... 178

Annex A (normative) **In-stream structures ... 179**

Annex B (normative) **SVC, MVC, and MVD sample group and sub-track definitions ... 193**

Annex C (normative) **Temporal metadata support ... 213**

Annex D (normative) **File format toolsets and brands ... 221**

Annex E (normative) **Sub-parameters for the MIME type 'codecs' parameter ... 224**

Annex F (informative) **Unspecified nal_unit_type value management for sample entry types of AVC and HEVC ... 231**

Annex G (informative) **Examples of VVC base and subpicture tracks ... 233**