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

**Part 10:
Advanced video coding**

*Technologies de l'information — Codage des objets audiovisuels —
Partie 10: Codage visuel avancé*

Withhold

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.

Withdrawn



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2008

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	xiii
0 Introduction	xv
0.1 Prologue.....	xv
0.2 Purpose.....	xv
0.3 Applications.....	xv
0.4 Publication and versions of this specification.....	xv
0.5 Profiles and levels.....	xvi
0.6 Overview of the design characteristics.....	xvi
0.6.1 Predictive coding.....	xvii
0.6.2 Coding of progressive and interlaced video.....	xvii
0.6.3 Picture partitioning into macroblocks and smaller partitions.....	xvii
0.6.4 Spatial redundancy reduction.....	xviii
0.7 How to read this specification.....	xviii
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviations	9
5 Conventions	10
5.1 Arithmetic operators.....	10
5.2 Logical operators.....	10
5.3 Relational operators.....	11
5.4 Bit-wise operators.....	11
5.5 Assignment operators.....	11
5.6 Range notation.....	11
5.7 Mathematical functions.....	12
5.8 Variables, syntax elements, and tables.....	13
5.9 Text description of logical operations.....	13
5.10 Processes.....	14
6 Source, coded, decoded and output data formats, scanning processes, and neighbouring relationships	15
6.1 Bitstream formats.....	15
6.2 Source, decoded, and output picture formats.....	15
6.3 Spatial subdivision of pictures and slices.....	20
6.4 Inverse scanning processes and derivation processes for neighbours.....	21
6.4.1 Inverse macroblock scanning process.....	21
6.4.2 Inverse macroblock partition and sub-macroblock partition scanning process.....	22
6.4.3 Inverse 4x4 luma block scanning process.....	23
6.4.4 Inverse 4x4 Cb or Cr block scanning process for ChromaArrayType equal to 3.....	23
6.4.5 Inverse 8x8 luma block scanning process.....	24
6.4.6 Inverse 8x8 Cb or Cr block scanning process for ChromaArrayType equal to 3.....	24
6.4.7 Derivation process of the availability for macroblock addresses.....	24
6.4.8 Derivation process for neighbouring macroblock addresses and their availability.....	24
6.4.9 Derivation process for neighbouring macroblock addresses and their availability in MBAFF frames.....	25
6.4.10 Derivation processes for neighbouring macroblocks, blocks, and partitions.....	26
6.4.11 Derivation process for neighbouring locations.....	31
7 Syntax and semantics	34
7.1 Method of specifying syntax in tabular form.....	34
7.2 Specification of syntax functions, categories, and descriptors.....	35
7.3 Syntax in tabular form.....	36
7.3.1 NAL unit syntax.....	36
7.3.2 Raw byte sequence payloads and RBSP trailing bits syntax.....	37
7.3.3 Slice header syntax.....	46

7.3.4	Slice data syntax	50
7.3.5	Macroblock layer syntax	51
7.4	<i>Semantics</i>	58
7.4.1	NAL unit semantics	58
7.4.2	Raw byte sequence payloads and RBSP trailing bits semantics	67
7.4.3	Slice header semantics	81
7.4.4	Slice data semantics	91
7.4.5	Macroblock layer semantics	92
8	Decoding process	104
8.1	<i>NAL unit decoding process</i>	105
8.2	<i>Slice decoding process</i>	106
8.2.1	Decoding process for picture order count	106
8.2.2	Decoding process for macroblock to slice group map	110
8.2.3	Decoding process for slice data partitioning	114
8.2.4	Decoding process for reference picture lists construction	114
8.2.5	Decoded reference picture marking process	121
8.3	<i>Intra prediction process</i>	125
8.3.1	Intra_4x4 prediction process for luma samples	126
8.3.2	Intra_8x8 prediction process for luma samples	132
8.3.3	Intra_16x16 prediction process for luma samples	140
8.3.4	Intra prediction process for chroma samples	142
8.3.5	Sample construction process for I_PCM macroblocks	147
8.4	<i>Inter prediction process</i>	147
8.4.1	Derivation process for motion vector components and reference indices	150
8.4.2	Decoding process for Inter prediction samples	163
8.4.3	Derivation process for prediction weights	172
8.5	<i>Transform coefficient decoding process and picture construction process prior to deblocking filter process</i>	174
8.5.1	Specification of transform decoding process for 4x4 luma residual blocks	174
8.5.2	Specification of transform decoding process for luma samples of Intra_16x16 macroblock prediction mode	175
8.5.3	Specification of transform decoding process for 8x8 luma residual blocks	176
8.5.4	Specification of transform decoding process for chroma samples	177
8.5.5	Specification of transform decoding process for chroma samples with ChromaArrayType equal to 3	179
8.5.6	Inverse scanning process for transform coefficients	179
8.5.7	Inverse scanning process for 8x8 transform coefficients	180
8.5.8	Derivation process for the chroma quantisation parameters and scaling function	182
8.5.9	Scaling and transformation process for DC transform coefficients for Intra_16x16 macroblock type	184
8.5.10	Scaling and transformation process for chroma DC transform coefficients	185
8.5.11	Scaling and transformation process for residual 4x4 blocks	186
8.5.12	Scaling and transformation process for residual 8x8 blocks	189
8.5.13	Picture construction process prior to deblocking filter process	193
8.5.14	Intra residual transform-bypass decoding process	194
8.6	<i>Decoding process for P macroblocks in SP slices or SI macroblocks</i>	195
8.6.1	SP decoding process for non-switching pictures	195
8.6.2	SP and SI slice decoding process for switching pictures	198
8.7	<i>Deblocking filter process</i>	200
8.7.1	Filtering process for block edges	204
8.7.2	Filtering process for a set of samples across a horizontal or vertical block edge	206
9	Parsing process	212
9.1	<i>Parsing process for Exp-Golomb codes</i>	212
9.1.1	Mapping process for signed Exp-Golomb codes	214
9.1.2	Mapping process for coded block pattern	214
9.2	<i>CAVLC parsing process for transform coefficient levels</i>	217
9.2.1	Parsing process for total number of transform coefficient levels and trailing ones	217
9.2.2	Parsing process for level information	221
9.2.3	Parsing process for run information	222
9.2.4	Combining level and run information	225
9.3	<i>CABAC parsing process for slice data</i>	225
9.3.1	Initialisation process	227

9.3.2	Binarization process	249
9.3.3	Decoding process flow	257
9.3.4	Arithmetic encoding process (informative).....	278
Annex A Profiles and levels		285
A.1	<i>Requirements on video decoder capability</i>	285
A.2	<i>Profiles</i>	285
A.2.1	Baseline profile	285
A.2.2	Main profile.....	286
A.2.3	Extended profile	286
A.2.4	High profile	286
A.2.5	High 10 profile	287
A.2.6	High 4:2:2 profile	287
A.2.7	High 4:4:4 Predictive profile	288
A.2.8	High 10 Intra profile.....	288
A.2.9	High 4:2:2 Intra profile	289
A.2.10	High 4:4:4 Intra profile	289
A.2.11	CAVLC 4:4:4 Intra profile	290
A.3	<i>Levels</i>	290
A.3.1	Level limits common to the Baseline, Main, and Extended profiles	290
A.3.2	Level limits common to the High, High 10, High 4:2:2, High 4:4:4 Predictive, High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, and CAVLC 4:4:4 Intra profiles	292
A.3.3	Profile-specific level limits	293
A.3.4	Effect of level limits on frame rate (informative).....	298
Annex B Byte stream format.....		301
B.1	<i>Byte stream NAL unit syntax and semantics</i>	301
B.1.1	Byte stream NAL unit syntax	301
B.1.2	Byte stream NAL unit semantics	301
B.2	<i>Byte stream NAL unit decoding process</i>	302
B.3	<i>Decoder byte-alignment recovery (informative)</i>	302
Annex C Hypothetical reference decoder		304
C.1	<i>Operation of coded picture buffer (CPB)</i>	306
C.1.1	Timing of bitstream arrival	306
C.1.2	Timing of coded picture removal	308
C.2	<i>Operation of the decoded picture buffer (DPB)</i>	308
C.2.1	Decoding of gaps in frame_num and storage of "non-existing" frames	308
C.2.2	Picture decoding and output	309
C.2.3	Removal of pictures from the DPB before possible insertion of the current picture	309
C.2.4	Current decoded picture marking and storage	310
C.3	<i>Bitstream conformance</i>	310
C.4	<i>Decoder conformance</i>	312
C.4.1	Operation of the output order DPB	313
C.4.2	Decoding of gaps in frame_num and storage of "non-existing" pictures	313
C.4.3	Picture decoding.....	313
C.4.4	Removal of pictures from the DPB before possible insertion of the current picture	313
C.4.5	Current decoded picture marking and storage	314
Annex D Supplemental enhancement information		316
D.1	<i>SEI payload syntax</i>	316
D.1.1	Buffering period SEI message syntax	318
D.1.2	Picture timing SEI message syntax	318
D.1.3	Pan-scan rectangle SEI message syntax	319
D.1.4	Filler payload SEI message syntax	319
D.1.5	User data registered by ITU-T Rec. T.35 SEI message syntax	320
D.1.6	User data unregistered SEI message syntax	320
D.1.7	Recovery point SEI message syntax	320
D.1.8	Decoded reference picture marking repetition SEI message syntax.....	321
D.1.9	Spare picture SEI message syntax	321
D.1.10	Scene information SEI message syntax.....	322
D.1.11	Sub-sequence information SEI message syntax	322
D.1.12	Sub-sequence layer characteristics SEI message syntax	322

D.1.13	Sub-sequence characteristics SEI message syntax.....	323
D.1.14	Full-frame freeze SEI message syntax.....	323
D.1.15	Full-frame freeze release SEI message syntax.....	323
D.1.16	Full-frame snapshot SEI message syntax.....	323
D.1.17	Progressive refinement segment start SEI message syntax.....	324
D.1.18	Progressive refinement segment end SEI message syntax.....	324
D.1.19	Motion-constrained slice group set SEI message syntax.....	324
D.1.20	Film grain characteristics SEI message syntax.....	325
D.1.21	Deblocking filter display preference SEI message syntax.....	325
D.1.22	Stereo video information SEI message syntax.....	326
D.1.23	Post-filter hint SEI message syntax.....	326
D.1.24	Tone mapping information SEI message syntax.....	327
D.1.25	Reserved SEI message syntax.....	327
D.2	<i>SEI payload semantics</i>	327
D.2.1	Buffering period SEI message semantics.....	327
D.2.2	Picture timing SEI message semantics.....	328
D.2.3	Pan-scan rectangle SEI message semantics.....	332
D.2.4	Filler payload SEI message semantics.....	333
D.2.5	User data registered by ITU-T Rec. T.35 SEI message semantics.....	333
D.2.6	User data unregistered SEI message semantics.....	333
D.2.7	Recovery point SEI message semantics.....	333
D.2.8	Decoded reference picture marking repetition SEI message semantics.....	335
D.2.9	Spare picture SEI message semantics.....	335
D.2.10	Scene information SEI message semantics.....	337
D.2.11	Sub-sequence information SEI message semantics.....	339
D.2.12	Sub-sequence layer characteristics SEI message semantics.....	340
D.2.13	Sub-sequence characteristics SEI message semantics.....	341
D.2.14	Full-frame freeze SEI message semantics.....	342
D.2.15	Full-frame freeze release SEI message semantics.....	343
D.2.16	Full-frame snapshot SEI message semantics.....	343
D.2.17	Progressive refinement segment start SEI message semantics.....	343
D.2.18	Progressive refinement segment end SEI message semantics.....	344
D.2.19	Motion-constrained slice group set SEI message semantics.....	344
D.2.20	Film grain characteristics SEI message semantics.....	345
D.2.21	Deblocking filter display preference SEI message semantics.....	350
D.2.22	Stereo video information SEI message semantics.....	352
D.2.23	Post-filter hint SEI message semantics.....	353
D.2.24	Tone mapping information SEI message semantics.....	354
D.2.25	Reserved SEI message semantics.....	355
Annex E	Video usability information	356
E.1	<i>VUI syntax</i>	357
E.1.1	VUI parameters syntax.....	357
E.1.2	HRD parameters syntax.....	358
E.2	<i>VUI semantics</i>	358
E.2.1	VUI parameters semantics.....	358
E.2.2	HRD parameters semantics.....	371
Annex F	Patent Rights	373
Annex G	Scalable video coding	375
G.1	<i>Scope</i>	375
G.2	<i>Normative References</i>	375
G.3	<i>Definitions</i>	375
G.4	<i>Abbreviations</i>	379
G.5	<i>Conventions</i>	379
G.6	<i>Source, coded, decoded and output data formats, scanning processes, neighbouring and reference layer relationships</i>	379
G.6.1	Derivation process for reference layer macroblocks.....	379
G.6.2	Derivation process for reference layer partitions.....	383
G.6.3	Derivation process for reference layer sample locations in resampling.....	384
G.7	<i>Syntax and semantics</i>	386

G.7.1	Method of specifying syntax in tabular form	386
G.7.2	Specification of syntax functions, categories, and descriptors	386
G.7.3	Syntax in tabular form	386
G.7.4	Semantics	397
G.7.4.1	NAL unit semantics	398
G.8	<i>SVC decoding process</i>	423
G.8.1	SVC initialization and decoding processes	424
G.8.2	SVC reference picture lists construction and decoded reference picture marking process	441
G.8.3	SVC intra decoding processes	450
G.8.4	SVC Inter prediction process	459
G.8.5	SVC transform coefficient decoding and sample array construction processes	470
G.8.6	Resampling processes for prediction data, intra samples, and residual samples	488
G.8.7	SVC deblocking filter processes	514
G.8.8	Specification of bitstream subsets	525
G.9	<i>Parsing process</i>	526
G.9.1	Alternative parsing process for coded block pattern	526
G.9.2	Alternative CAVLC parsing process for transform coefficient levels	527
G.9.3	Alternative CABAC parsing process for slice data in scalable extension	531
G.10	<i>Profiles and levels</i>	534
G.10.1	Profiles	534
G.10.2	Levels	537
G.11	<i>Byte stream format</i>	541
G.12	<i>Hypothetical reference decoder</i>	541
G.13	<i>Supplemental enhancement information</i>	541
G.13.1	SEI payload syntax	542
G.13.2	SEI payload semantics	549
G.14	<i>SVC video usability information extension</i>	565
G.14.1	SVC VUI parameters extension syntax	565
G.14.2	SVC VUI parameters extension semantics	565

LIST OF FIGURES

Figure 6-1	– Nominal vertical and horizontal locations of 4:2:0 luma and chroma samples in a frame	16
Figure 6-2	– Nominal vertical and horizontal sampling locations of 4:2:0 samples in top and bottom fields	17
Figure 6-3	– Nominal vertical and horizontal locations of 4:2:2 luma and chroma samples in a frame	17
Figure 6-4	– Nominal vertical and horizontal sampling locations of 4:2:2 samples top and bottom fields	18
Figure 6-5	– Nominal vertical and horizontal locations of 4:4:4 luma and chroma samples in a frame	18
Figure 6-6	– Nominal vertical and horizontal sampling locations of 4:4:4 samples top and bottom fields	19
Figure 6-7	– A picture with 11 by 9 macroblocks that is partitioned into two slices	20
Figure 6-8	– Partitioning of the decoded frame into macroblock pairs	20
Figure 6-9	– Macroblock partitions, sub-macroblock partitions, macroblock partition scans, and sub-macroblock partition scans	22
Figure 6-10	– Scan for 4x4 luma blocks	23
Figure 6-11	– Scan for 8x8 luma blocks	23
Figure 6-12	– Neighbouring macroblocks for a given macroblock	24
Figure 6-13	– Neighbouring macroblocks for a given macroblock in MBAFF frames	25
Figure 6-14	– Determination of the neighbouring macroblock, blocks, and partitions (informative)	26

Figure 7-1 – Structure of an access unit not containing any NAL units with nal_unit_type equal to 0, 7, 8, or in the range of 12 to 18, inclusive, or in the range of 20 to 31, inclusive.....	61
Figure 8-1 – Intra_4x4 prediction mode directions (informative).....	122
Figure 8-2 – Example for temporal direct-mode motion vector inference (informative).....	152
Figure 8-3 – Directional segmentation prediction (informative).....	153
Figure 8-4 – Integer samples (shaded blocks with upper-case letters) and fractional sample positions (un-shaded blocks with lower-case letters) for quarter sample luma interpolation.....	159
Figure 8-5 – Fractional sample position dependent variables in chroma interpolation and surrounding integer position samples A, B, C, and D.....	161
Figure 8-6 – Assignment of the indices of dcY to luma4x4BlkIdx.....	167
Figure 8-7 – Assignment of the indices of dcC to chroma4x4BlkIdx: (a) ChromaArrayType equal to 1, (b) ChromaArrayType equal to 2.....	169
Figure 8-8 – 4x4 block scans. (a) Zig-zag scan. (b) Field scan (informative).....	171
Figure 8-9 – 8x8 block scans. (a) 8x8 zig-zag scan. (b) 8x8 field scan (informative).....	172
Figure 8-10 – Boundaries in a macroblock to be filtered.....	191
Figure 8-11 – Convention for describing samples across a 4x4 block horizontal or vertical boundary.....	195
Figure 9-1 – Illustration of CABAC parsing process for a syntax element SE (informative).....	216
Figure 9-2 – Overview of the arithmetic decoding process for a single bin (informative).....	261
Figure 9-3 – Flowchart for decoding a decision.....	262
Figure 9-4 – Flowchart of renormalization.....	264
Figure 9-5 – Flowchart of bypass decoding process.....	265
Figure 9-6 – Flowchart of decoding a decision before termination.....	266
Figure 9-7 – Flowchart for encoding a decision.....	267
Figure 9-8 – Flowchart of renormalization in the encoder.....	268
Figure 9-9 – Flowchart of PutBin(B).....	268
Figure 9-10 – Flowchart of encoding bypass.....	269
Figure 9-11 – Flowchart of encoding a decision before termination.....	270
Figure 9-12 – Flowchart of flushing at termination.....	270
Figure C-1 – Structure of byte streams and NAL unit streams for HRD conformance checks.....	290
Figure C-2 – HRD buffer model.....	291
Figure E-1 – Location of chroma samples for top and bottom fields as a function of chroma_sample_loc_type_top_field and chroma_sample_loc_type_bottom_field.....	352

LIST OF TABLES

Table 6-1 – SubWidthC, and SubHeightC values derived from chroma_format_idc and separate_colour_plane_flag.....	15
Table 6-2 – Specification of input and output assignments for subclauses 6.4.10.1 to 6.4.10.7.....	26
Table 6-3 – Specification of mbAddrN.....	31
Table 6-4 – Specification of mbAddrN and yM.....	33
Table 7-1 – NAL unit type codes, syntax element categories, and NAL unit type classes.....	59

Table 7-2 – Assignment of mnemonic names to scaling list indices and specification of fall-back rule	70
Table 7-3 – Specification of default scaling lists Default_4x4_Intra and Default_4x4_Inter	70
Table 7-4 – Specification of default scaling lists Default_8x8_Intra and Default_8x8_Inter	71
Table 7-5 – Meaning of primary_pic_type	78
Table 7-6 – Name association to slice_type	81
Table 7-7 – reordering_of_pic_nums_idc operations for reordering of reference picture lists	87
Table 7-8 – Interpretation of adaptive_ref_pic_marking_mode_flag	89
Table 7-9 – Memory management control operation (memory_management_control_operation) values	90
Table 7-10 – Allowed collective macroblock types for slice_type	92
Table 7-11 – Macroblock types for I slices	93
Table 7-12 – Macroblock type with value 0 for SI slices	94
Table 7-13 – Macroblock type values 0 to 4 for P and SP slices	95
Table 7-14 – Macroblock type values 0 to 22 for B slices	96
Table 7-15 – Specification of CodedBlockPatternChroma values	98
Table 7-16 – Relationship between intra_chroma_pred_mode and spatial prediction modes	99
Table 7-17 – Sub-macroblock types in P macroblocks	100
Table 7-18 – Sub-macroblock types in B macroblocks	101
Table 8-1 – Refined slice group map type	111
Table 8-2 – Specification of Intra4x4PredMode[luma4x4BlkIdx] and associated names	127
Table 8-3 – Specification of Intra8x8PredMode[luma8x8BlkIdx] and associated names	133
Table 8-4 – Specification of Intra16x16PredMode and associated names	141
Table 8-5 – Specification of Intra chroma prediction modes and associated names	143
Table 8-6 – Specification of the variable colPic	152
Table 8-7 – Specification of PicCodingStruct(X)	153
Table 8-8 – Specification of mbAddrCol, yM, and vertMvScale	154
Table 8-9 – Assignment of prediction utilization flags	156
Table 8-10 – Derivation of the vertical component of the chroma vector in field coding mode	162
Table 8-11 – Differential full-sample luma locations	167
Table 8-12 – Assignment of the luma prediction sample predPartLX _i [x _L , y _L]	169
Table 8-13 – Specification of mapping of idx to c _{ij} for zig-zag and field scan	180
Table 8-14 – Specification of mapping of idx to c _{ij} for 8x8 zig-zag and 8x8 field scan	181
Table 8-15 – Specification of QP _C as a function of qP _I	182
Table 8-16 – Derivation of offset dependent threshold variables α' and β' from indexA and indexB	209
Table 8-17 – Value of variable t' _{C0} as a function of indexA and bS	211
Table 9-1 – Bit strings with “prefix” and “suffix” bits and assignment to codeNum ranges (informative)	213
Table 9-2 – Exp-Golomb bit strings and codeNum in explicit form and used as ue(v) (informative)	213
Table 9-3 – Assignment of syntax element to codeNum for signed Exp-Golomb coded syntax elements se(v)	214
Table 9-4 – Assignment of codeNum to values of coded_block_pattern for macroblock prediction modes	215

Table 9-5 – coeff_token mapping to TotalCoeff(coeff_token) and TrailingOnes(coeff_token)	219
Table 9-6 – Codeword table for level_prefix (informative)	222
Table 9-7 – total_zeros tables for 4x4 blocks with tzVlcIndex 1 to 7	223
Table 9-8 – total_zeros tables for 4x4 blocks with tzVlcIndex 8 to 15	224
Table 9-9 – total_zeros tables for chroma DC 2x2 and 2x4 blocks	224
Table 9-10 – Tables for run_before	225
Table 9-11 – Association of ctxIdx and syntax elements for each slice type in the initialisation process	228
Table 9-12 – Values of variables m and n for ctxIdx from 0 to 10	229
Table 9-13 – Values of variables m and n for ctxIdx from 11 to 23	229
Table 9-14 – Values of variables m and n for ctxIdx from 24 to 39	230
Table 9-15 – Values of variables m and n for ctxIdx from 40 to 53	230
Table 9-16 – Values of variables m and n for ctxIdx from 54 to 59, and 399 to 401	230
Table 9-17 – Values of variables m and n for ctxIdx from 60 to 69	231
Table 9-18 – Values of variables m and n for ctxIdx from 70 to 104	231
Table 9-19 – Values of variables m and n for ctxIdx from 105 to 165	232
Table 9-20 – Values of variables m and n for ctxIdx from 166 to 226	233
Table 9-21 – Values of variables m and n for ctxIdx from 227 to 275	234
Table 9-22 – Values of variables m and n for ctxIdx from 277 to 337	235
Table 9-23 – Values of variables m and n for ctxIdx from 338 to 398	236
Table 9-24 – Values of variables m and n for ctxIdx from 402 to 459	237
Table 9-25 – Values of variables m and n for ctxIdx from 460 to 483	238
Table 9-26 – Values of variables m and n for ctxIdx from 484 to 571	239
Table 9-27 – Values of variables m and n for ctxIdx from 572 to 659	241
Table 9-28 – Values of variables m and n for ctxIdx from 660 to 717	242
Table 9-29 – Values of variables m and n for ctxIdx from 718 to 775	244
Table 9-30 – Values of variables m and n for ctxIdx from 776 to 863	245
Table 9-31 – Values of variables m and n for ctxIdx from 864 to 951	246
Table 9-32 – Values of variables m and n for ctxIdx from 952 to 1011	248
Table 9-33 – Values of variables m and n for ctxIdx from 1012 to 1023	249
Table 9-34 – Syntax elements and associated types of binarization, maxBinIdxCtx, and ctxIdxOffset	250
Table 9-35 – Bin string of the unary binarization (informative)	253
Table 9-36 – Binarization for macroblock types in I slices	255
Table 9-37 – Binarization for macroblock types in P, SP, and B slices	256
Table 9-38 – Binarization for sub-macroblock types in P, SP, and B slices	257
Table 9-39 – Assignment of ctxIdxInc to binIdx for all ctxIdxOffset values except those related to the syntax elements coded_block_flag, significant_coeff_flag, last_significant_coeff_flag, and coeff_abs_level_minus1	259
Table 9-40 – Assignment of ctxIdxBlockCatOffset to ctxBlockCat for syntax elements coded_block_flag, significant_coeff_flag, last_significant_coeff_flag, and coeff_abs_level_minus1	260

Table 9-41 – Specification of ctxIdxInc for specific values of ctxIdxOffset and binIdx	269
Table 9-42 – Specification of ctxBlockCat for the different blocks.....	269
Table 9-43 – Mapping of scanning position to ctxIdxInc for ctxBlockCat == 5, 9, or 13	270
Table 9-44 – Specification of rangeTabLPS depending on pStateIdx and qCodIRangeIdx	275
Table 9-45 – State transition table	276
Table A-1 – Level limits	292
Table A-2 – Specification of cpbBrVclFactor and cpbBrNalFactor	295
Table A-3 – Baseline profile level limits	295
Table A-4 – Main, High, High 10, High 4:2:2, High 4:4:4 Predictive, High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, and CAVLC 4:4:4 Intra profile level limits	296
Table A-5 – Extended profile level limits	297
Table A-6 – Maximum frame rates (frames per second) for some example frame sizes	298
Table D-1 – Interpretation of pic_struct.....	329
Table D-2 – Mapping of ct_type to source picture scan	330
Table D-3 – Definition of counting_type values.....	330
Table D-4 – scene_transition_type values	337
Table D-5 – model_id values	345
Table D-6 – blending_mode_id values	346
Table D-7 – filter_hint_type values	353
Table E-1 – Meaning of sample aspect ratio indicator.....	359
Table E-2 – Meaning of video_format.....	360
Table E-3 – Colour primaries.....	361
Table E-4 – Transfer characteristics.....	362
Table E-5 – Matrix coefficients	366
Table E-6 – Divisor for computation of $\Delta_{fi,dpb}(n)$	368
Table F-1 – Organizations providing patent rights licensing notices	373
Table G-1 – Name association to slice_type for NAL units with nal_unit_type equal to 20.....	411
Table G-2 – Memory management control operation (memory_management_control_operation) values.....	419
Table G-3 – Allowed collective macroblock types for slice_type.	421
Table G-4 – Macroblock types for EI slices.....	421
Table G-5 – Scale values cS for transform coefficient level scaling.....	479
Table G-6 – Macroblock type predictors mbTypeILPred	497
Table G-7 – Sub-macroblock type predictors subMbTypeILPred[mbPartIdx].....	497
Table G-8 – 16-phase luma interpolation filter for resampling in Intra_Base prediction	505
Table G-9 – Mapping of (nX, nY) to coeffTokenIdx and vice versa	528
Table G-10 - Association of ctxIdx and syntax elements for each slice type in the initialisation process.....	532
Table G-11 - Values of variables m and n for ctxIdx from 1024 to 1026	532
Table G-12 - Values of variables m and n for ctxIdx from 1027 to 1029	532

Table G-13 - Syntax elements and associated types of binarization, maxBinIdxCtx, and ctxIdxOffset	533
Table G-14 - Assignment of ctxIdxInc to binIdx for the ctxIdxOffset values related to the syntax elements base_mode_flag and residual_prediction_flag	533
Table G-15 – Scalable Baseline profile level limits	541
Table G-16 – Specification of cpbBrVclFactor and cpbBrNalFactor.....	541

Withdrawn

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 14496-10 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and information* in collaboration with ITU-T.

This part of ISO/IEC 14496 is technically aligned with ITU-T Rec. H.264 but is not published as identical text.

This fourth edition cancels and replaces the third edition 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* [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 (LAsEeR) 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*

Withdrawn

0 Introduction

This clause does not form an integral part of this International Standard.

0.1 Prologue

This subclause does not form an integral part of this International Standard.

As the costs for both processing power and memory have reduced, network support for coded video data has diversified, and advances in video coding technology have progressed, the need has arisen for an industry standard for compressed video representation with substantially increased coding efficiency and enhanced robustness to network environments. Toward these ends the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) formed a Joint Video Team (JVT) in 2001 for development of a new International Standard.

0.2 Purpose

This subclause does not form an integral part of this International Standard.

This International Standard was developed in response to the growing need for higher compression of moving pictures for various applications such as videoconferencing, digital storage media, television broadcasting, internet streaming, and communication. It is also designed to enable the use of the coded video representation in a flexible manner for a wide variety of network environments. The use of this International Standard allows motion video to be manipulated as a form of computer data and to be stored on various storage media, transmitted and received over existing and future networks and distributed on existing and future broadcasting channels.

0.3 Applications

This subclause does not form an integral part of this International Standard.

This International Standard is designed to cover a broad range of applications for video content including but not limited to the following:

CATV	Cable TV on optical networks, copper, etc.
DBS	Direct broadcast satellite video services
DSL	Digital subscriber line video services
DTTB	Digital terrestrial television broadcasting
ISM	Interactive storage media (optical disks, etc.)
MMM	Multimedia mailing
MSPN	Multimedia services over packet networks
RTC	Real-time conversational services (videoconferencing, videophone, etc.)
RVS	Remote video surveillance
SSM	Serial storage media (digital VTR, etc.)

0.4 Publication and versions of this specification

This subclause does not form an integral part of this International Standard.

This specification has been jointly developed by ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group. It is published as technically-aligned twin text in both organizations ITU-T and ISO/IEC.

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 1 refers to the first approved version of this International Standard.

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 2 refers to the integrated text containing the corrections specified in the first technical corrigendum.

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 3 refers to the integrated text containing both the first technical corrigendum (2004) and the first amendment, which is referred to as the "Fidelity range extensions".

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 4 refers to the integrated text containing the first technical corrigendum (2004), the first amendment (the "Fidelity range extensions"), and an additional technical corrigendum (2005).

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 5 refers to the integrated version 4 text with its specification of the High 4:4:4 profile removed.

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 6 refers to the integrated version 5 text after its amendment to support additional colour space indicators.

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 7 refers to the integrated version 6 text after its amendment to define five new profiles intended primarily for professional applications (the High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, CAVLC 4:4:4 Intra, and High 4:4:4 Predictive profiles) and two new types of supplemental enhancement information (SEI) messages (the post-filter hint SEI message and the tone mapping information SEI message).

ITU-T Rec. H.264 | ISO/IEC 14496-10 version 8 (the current specification) refers to the integrated version 7 text after its amendment to specify scalable video coding in three profiles (Scalable Baseline, Scalable High, and Scalable High Intra profiles).

0.5 Profiles and levels

This subclause does not form an integral part of this International Standard.

This International Standard is designed to be generic in the sense that it serves a wide range of applications, bit rates, resolutions, qualities, and services. Applications should cover, among other things, digital storage media, television broadcasting and real-time communications. In the course of creating this Specification, various requirements from typical applications have been considered, necessary algorithmic elements have been developed, and these have been integrated into a single syntax. Hence, this Specification will facilitate video data interchange among different applications.

Considering the practicality of implementing the full syntax of this Specification, however, a limited number of subsets of the syntax are also stipulated by means of "profiles" and "levels". These and other related terms are formally defined in clause 3.

A "profile" is a subset of the entire bitstream syntax that is specified by this International Standard. Within the bounds imposed by the syntax of a given profile it is still possible to require a very large variation in the performance of encoders and decoders depending upon the values taken by syntax elements in the bitstream such as the specified size of the decoded pictures. In many applications, it is currently neither practical nor economic to implement a decoder capable of dealing with all hypothetical uses of the syntax within a particular profile.

In order to deal with this problem, "levels" are specified within each profile. A level is a specified set of constraints imposed on values of the syntax elements in the bitstream. These constraints may be simple limits on values. Alternatively they may take the form of constraints on arithmetic combinations of values (e.g. picture width multiplied by picture height multiplied by number of pictures decoded per second).

Coded video content conforming to this International Standard uses a common syntax. In order to achieve a subset of the complete syntax, flags, parameters, and other syntax elements are included in the bitstream that signal the presence or absence of syntactic elements that occur later in the bitstream.

0.6 Overview of the design characteristics

This subclause does not form an integral part of this International Standard.

The coded representation specified in the syntax is designed to enable a high compression capability for a desired image quality. With the exception of the transform bypass mode of operation for lossless coding in the High 4:4:4 Intra, CAVLC 4:4:4 Intra, and High 4:4:4 Predictive profiles, and the I_PCM mode of operation in all profiles, the algorithm is typically not lossless, as the exact source sample values are typically not preserved through the encoding and decoding processes. A number of techniques may be used to achieve highly efficient compression. Encoding algorithms (not specified in this International Standard) may select between inter and intra coding for block-shaped regions of each picture. Inter coding uses motion vectors for block-based inter prediction to exploit temporal statistical dependencies between different pictures. Intra coding uses various spatial prediction modes to exploit spatial statistical dependencies in the source signal for a single picture. Motion vectors and intra prediction modes may be specified for a variety of block sizes in the picture. The prediction residual is then further compressed using a transform to remove spatial correlation inside the transform block before it is quantised, producing an irreversible process that typically discards less important visual information while forming a close approximation to the source samples. Finally, the motion vectors or intra

prediction modes are combined with the quantised transform coefficient information and encoded using either variable length codes or arithmetic coding.

Scalable video coding is specified in Annex G of this specification allowing the construction of bitstreams that contain sub-bitstreams that conform to this specification. For temporal bitstream scalability, i.e., the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, i.e. the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL units are removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, i.e., the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding.

Otherwise, the coding algorithm as described in the previous paragraph is used.

0.6.1 Predictive coding

This subclause does not form an integral part of this International Standard.

Because of the conflicting requirements of random access and highly efficient compression, two main coding types are specified. Intra coding is done without reference to other pictures. Intra coding may provide access points to the coded sequence where decoding can begin and continue correctly, but typically also shows only moderate compression efficiency. Inter coding (predictive or bi-predictive) is more efficient using inter prediction of each block of sample values from some previously decoded picture selected by the encoder. In contrast to some other video coding standards, pictures coded using bi-predictive inter prediction may also be used as references for inter coding of other pictures.

The application of the three coding types to pictures in a sequence is flexible, and the order of the decoding process is generally not the same as the order of the source picture capture process in the encoder or the output order from the decoder for display. The choice is left to the encoder and will depend on the requirements of the application. The decoding order is specified such that the decoding of pictures that use inter-picture prediction follows later in decoding order than other pictures that are referenced in the decoding process.

0.6.2 Coding of progressive and interlaced video

This subclause does not form an integral part of this International Standard.

This International Standard specifies a syntax and decoding process for video that originated in either progressive-scan or interlaced-scan form, which may be mixed together in the same sequence. The two fields of an interlaced frame are separated in capture time while the two fields of a progressive frame share the same capture time. Each field may be coded separately or the two fields may be coded together as a frame. Progressive frames are typically coded as a frame. For interlaced video, the encoder can choose between frame coding and field coding. Frame coding or field coding can be adaptively selected on a picture-by-picture basis and also on a more localized basis within a coded frame. Frame coding is typically preferred when the video scene contains significant detail with limited motion. Field coding typically works better when there is fast picture-to-picture motion.

0.6.3 Picture partitioning into macroblocks and smaller partitions

This subclause does not form an integral part of this International Standard.

As in previous video coding Recommendations and International Standards, a macroblock, consisting of a 16x16 block of luma samples and two corresponding blocks of chroma samples, is used as the basic processing unit of the video decoding process.

A macroblock can be further partitioned for inter prediction. The selection of the size of inter prediction partitions is a result of a trade-off between the coding gain provided by using motion compensation with smaller blocks and the quantity of data needed to represent the data for motion compensation. In this International Standard the inter prediction process can form segmentations for motion representation as small as 4x4 luma samples in size, using motion vector accuracy of one-quarter of the luma sample grid spacing displacement. The process for inter prediction of a sample block can also involve the selection of the picture to be used as the reference picture from a number of stored previously-decoded pictures. Motion vectors are encoded differentially with respect to predicted values formed from nearby encoded motion vectors.

Typically, the encoder calculates appropriate motion vectors and other data elements represented in the video data stream. This motion estimation process in the encoder and the selection of whether to use inter prediction for the representation of each region of the video content is not specified in this International Standard.

0.6.4 Spatial redundancy reduction

This subclause does not form an integral part of this International Standard.

Both source pictures and prediction residuals have high spatial redundancy. This International Standard is based on the use of a block-based transform method for spatial redundancy removal. After inter prediction from previously-decoded samples in other pictures or spatial-based prediction from previously-decoded samples within the current picture, the resulting prediction residual is split into 4x4 blocks. These are converted into the transform domain where they are quantised. After quantisation many of the transform coefficients are zero or have low amplitude and can thus be represented with a small amount of encoded data. The processes of transformation and quantisation in the encoder are not specified in this International Standard.

0.7 How to read this specification

This subclause does not form an integral part of this International Standard.

It is suggested that the reader starts with clause 1 (Scope) and moves on to clause 3 (Definitions). Clause 6 should be read for the geometrical relationship of the source, input, and output of the decoder. Clause 7 (Syntax and semantics) specifies the order to parse syntax elements from the bitstream. See subclauses 7.1-7.3 for syntactical order and see subclause 7.4 for semantics; i.e., the scope, restrictions, and conditions that are imposed on the syntax elements. The actual parsing for most syntax elements is specified in clause 9 (Parsing process). Finally, clause 8 (Decoding process) specifies how the syntax elements are mapped into decoded samples. Throughout reading this specification, the reader should refer to clauses 2 (Normative references), 4 (Abbreviations), and 5 (Conventions) as needed. Annexes A through E also form an integral part of this International Standard.

Annex A specifies eleven profiles (Baseline, Main, Extended, High, High 10, High 4:2:2, High 4:4:4 Predictive, High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, and CAVLC 4:4:4 Intra), each being tailored to certain application domains, and defines the so-called levels of the profiles. Annex B specifies syntax and semantics of a byte stream format for delivery of coded video as an ordered stream of bytes. Annex C specifies the hypothetical reference decoder and its use to check bitstream and decoder conformance. Annex D specifies syntax and semantics for supplemental enhancement information message payloads. Finally, Annex E specifies syntax and semantics of the video usability information parameters of the sequence parameter set.

Annex G specifies scalable video coding (SVC). The reader is referred to Annex G for the entire decoding process for SVC, which is specified there with references being made to clauses 2-9 and Annexes A-E. Subclause G.10 specifies three profiles for SVC (Scalable Baseline, Scalable High, and Scalable High Intra).

Throughout this specification, statements appearing with the preamble "NOTE -" are informative and are not an integral part of this International Standard.

Information technology — Coding of audio-visual objects —

Part 10:

Advanced video coding

1 Scope

This part of ISO 14496 specifies advanced video coding for the coding of audio-visual objects.

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.

- ITU-T Recommendation T.35 (2000), *Procedure for the allocation of ITU-T defined codes for non-standard facilities*.
- ISO/IEC 11578:1996, Annex A, *Universal Unique Identifier*.
- ISO/CIE 10527:1991, *Colorimetric Observers*.