This is a preview - click here to buy the full publication

INTERNATIONAL STANDARD

# ISO/IEC 23090-3

First edition
2021-02

# Information technology — Coded representation of immersive media —

## Part 3: Versatile video coding

*Technologies de l'information — Représentation codée de média immersifs —*

*Partie 3: Codage vidéo polyvalent*

Withdrawn

Reference number
ISO/IEC 23090-3:2021(E)

© ISO/IEC 2021

This is a preview - click here to buy the full publication

withdrawn

**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2021

All rights reserved. Unless otherwise specified, or required in the context of its implementation, 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
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland



© ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication



# Contents

Foreword ... vi
Introduction ... vii
1 Scope ... 1
2 Normative references ... 1
3 Terms and definitions ... 1
4 Abbreviated terms ... 13
5 Conventions ... 16
5.1 General ... 16
5.2 Arithmetic operators ... 16
5.3 Logical operators ... 16
5.4 Relational operators ... 17
5.5 Bit-wise operators ... 17
5.6 Assignment operators ... 17
5.7 Range notation ... 17
5.8 Mathematical functions ... 17
5.9 Order of operation precedence ... 18
5.10 Variables, syntax elements and tables ... 19
5.11 Text description of logical operations ... 20
5.12 Processes ... 21
6 Bitstream and picture formats, partitionings, scanning processes and neighbouring relationships ... 21
6.1 Bitstream formats ... 21
6.2 Source, decoded and output picture formats ... 21
6.3 Partitioning of pictures, subpictures, slices, tiles, and CTUs ... 23
6.3.1 Partitioning of pictures into subpictures, slices, and tiles ... 23
6.3.2 Block, quadtree and multi-type tree structures ... 25
6.3.3 Spatial or component-wise partitionings ... 26
6.4 Availability processes ... 27
6.4.1 Allowed quad split process ... 27
6.4.2 Allowed binary split process ... 27
6.4.3 Allowed ternary split process ... 29
6.4.4 Derivation process for neighbouring block availability ... 30
6.5 Scanning processes ... 30
6.5.1 CTB raster scanning, tile scanning, and subpicture scanning processes ... 30
6.5.2 Up-right diagonal scan order array initialization process ... 34
6.5.3 Horizontal and vertical traverse scan order array initialization process ... 35
7 Syntax and semantics ... 35
7.1 Method of specifying syntax in tabular form ... 35
7.2 Specification of syntax functions and descriptors ... 36
7.3 Syntax in tabular form ... 38
7.3.1 NAL unit syntax ... 38
7.3.2 Raw byte sequence payloads, trailing bits and byte alignment syntax ... 38
7.3.3 Profile, tier, and level syntax ... 57
7.3.4 DPB parameters syntax ... 60
7.3.5 Timing and HRD parameters syntax ... 60
7.3.6 Supplemental enhancement information message syntax ... 61
7.3.7 Slice header syntax ... 62
7.3.8 Weighted prediction parameters syntax ... 64
7.3.9 Reference picture lists syntax ... 65
7.3.10 Reference picture list structure syntax ... 66
7.3.11 Slice data syntax ... 66
7.4 Semantics ... 88
7.4.1 General ... 88
7.4.2 NAL unit semantics ... 88

Withdrawn

© ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication

7.4.3 Raw byte sequence payloads, trailing bits and byte alignment semantics ..... 95
7.4.4 Profile, tier, and level semantics ..... 141
7.4.5 DPB parameters semantics ..... 146
7.4.6 Timing and HRD parameters semantics ..... 146
7.4.7 Supplemental enhancement information message semantics ..... 150
7.4.8 Slice header semantics ..... 150
7.4.9 Weighted prediction parameters semantics ..... 158
7.4.10 Reference picture lists semantics ..... 159
7.4.11 Reference picture list structure semantics ..... 160
7.4.12 Slice data semantics ..... 161
8 Decoding process ..... 183
8.1 General decoding process ..... 183
8.1.1 General ..... 183
8.1.2 Decoding process for a coded picture ..... 184
8.2 NAL unit decoding process ..... 185
8.3 Slice decoding process ..... 185
8.3.1 Decoding process for picture order count ..... 185
8.3.2 Decoding process for reference picture lists construction ..... 187
8.3.3 Decoding process for reference picture marking ..... 191
8.3.4 Decoding process for generating unavailable reference pictures ..... 192
8.3.5 Decoding process for symmetric motion vector difference reference indices ..... 192
8.3.6 Decoding process for collocated picture and no backward prediction ..... 193
8.4 Decoding process for coding units coded in intra prediction mode ..... 194
8.4.1 General decoding process for coding units coded in intra prediction mode ..... 194
8.4.2 Derivation process for luma intra prediction mode ..... 195
8.4.3 Derivation process for chroma intra prediction mode ..... 198
8.4.4 Cross-component chroma intra prediction mode checking process ..... 199
8.4.5 Decoding process for intra blocks ..... 200
8.5 Decoding process for coding units coded in inter prediction mode ..... 231
8.5.1 General decoding process for coding units coded in inter prediction mode ..... 231
8.5.2 Derivation process for motion vector components and reference indices ..... 235
8.5.3 Decoder-side motion vector refinement process ..... 254
8.5.4 Derivation process for geometric partitioning mode motion vector components and reference indices ..... 259
8.5.5 Derivation process for subblock motion vector components and reference indices ..... 260
8.5.6 Decoding process for inter blocks ..... 285
8.5.7 Decoding process for geometric partitioning mode inter blocks ..... 307
8.5.8 Decoding process for the residual signal of coding blocks coded in inter prediction mode ..... 313
8.5.9 Decoding process for the reconstructed signal of chroma coding blocks coded in inter prediction mode ..... 314
8.6 Decoding process for coding units coded in IBC prediction mode ..... 316
8.6.1 General decoding process for coding units coded in IBC prediction mode ..... 316
8.6.2 Derivation process for block vector components for IBC blocks ..... 317
8.6.3 Decoding process for IBC blocks ..... 321
8.7 Scaling, transformation and array construction process ..... 322
8.7.1 Derivation process for quantization parameters ..... 322
8.7.2 Scaling and transformation process ..... 324
8.7.3 Scaling process for transform coefficients ..... 325
8.7.4 Transformation process for scaled transform coefficients ..... 327
8.7.5 Picture reconstruction process ..... 347
8.8 In-loop filter process ..... 350
8.8.1 General ..... 350
8.8.2 Picture inverse mapping process for luma samples ..... 350
8.8.3 Deblocking filter process ..... 351
8.8.4 Sample adaptive offset process ..... 377
8.8.5 Adaptive loop filter process ..... 379
9 Parsing process ..... 391
9.1 General ..... 391
9.2 Parsing process for k-th order Exp-Golomb codes ..... 391
9.2.1 General ..... 391
9.2.2 Mapping process for signed Exp-Golomb codes ..... 392

Withdrawn

 © ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication



9.3 CABAC parsing process for slice data ..... 393
9.3.1 General ..... 393
9.3.2 Initialization process ..... 394
9.3.3 Binarization process ..... 418
9.3.4 Decoding process flow ..... 427
Annex A (normative) Profiles, tiers and levels ..... 444
Annex B (normative) Byte stream format ..... 456
Annex C (normative) Hypothetical reference decoder ..... 458
Annex D (normative) Supplemental enhancement information and use of SEI and VUI ..... 480
Bibliography ..... 504

Withdrawn

© ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication

# 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.

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 document 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 and IEC 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) or the IEC list of patent declarations received (see http://patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of 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 www.iso.org/iso/foreword.html.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*, in collaboration with ITU-T. The technically identical text is published as ITU-T H.266 (08/2020).

A list of all parts in the ISO/IEC 23090 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Withdrawn

 © ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication



# Introduction

## Purpose

This document specifies a video coding technology known as versatile video coding. It has been designed with two primary goals. The first of these is to specify a video coding technology with a compression capability that is substantially beyond that of the prior generations of such standards, and the second is for this technology to be highly versatile for effective use in a broader range of applications than that addressed by prior standards. Some key application areas for the use of this document particularly include ultra-high-definition video (e.g., with 3840×2160 or 7620×4320 picture resolution and bit depth of 10 bits as specified in Rec. ITU-R BT.2100), video with a high dynamic range and wide colour gamut (e.g., with the perceptual quantization or hybrid log-gamma transfer characteristics specified in Rec. ITU-R BT.2100), and video for immersive media applications such as 360° omnidirectional video projected using a common projection format such as the equirectangular or cubemap projection formats, in addition to the applications that have commonly been addressed by prior video coding standards.

## Profiles, tiers, and levels

This document is designed to be versatile in the sense that it serves a wide range of applications, bit rates, resolutions, qualities, and services. Applications include, but are not limited to, video coding for digital storage media, television broadcasting, video streaming services, real-time communication. In the course of creating this document, 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 document is designed to facilitate video data interchange among different applications.

Considering the practicality of implementing the full syntax of this document, however, a limited number of subsets of the syntax are also stipulated by means of "profiles", "tiers", 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 in this document. 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 economical 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, "tiers" and "levels" are specified within each profile. A level of a tier is a specified set of constraints imposed on values of the syntax elements in the bitstream. Some of these constraints are expressed as simple limits on values, while others 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). A level specified for a lower tier is more constrained than a level specified for a higher tier.

Coded video content conforming to this document 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.

## Encoding process, decoding process, and use of VUI parameters and SEI messages

Any encoding process that produces bitstream data that conforms to the specified bitstream syntax format requirements of this document is considered to be in conformance with the requirements of this document. The decoding process is specified such that all decoders that conform to a specified combination of capabilities known as the profile, tier, and level will produce numerically identical cropped decoded output pictures when invoking the decoding process associated with that profile for a bitstream conforming to that profile, tier and level. Any decoding process that produces identical cropped decoded output pictures to those produced by the process described herein (with the correct output order or output timing, as specified) is considered to be in conformance with the requirements of this document.

Rec. ITU-T H.274 | ISO/IEC 23002-7 specifies the syntax and semantics of the video usability information (VUI) parameters and supplemental enhancement information (SEI) messages that do not affect the conformance specifications in Annex C. These VUI parameters and SEI messages may be used together with this document.

## Versions of this document

This is the first edition of this document.

Withdrawn

© ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication

## Overview of the design characteristics

The coded representation specified in the syntax is designed to enable a high compression capability for a desired image or video quality. The algorithm is typically not mathematically lossless, as the exact source sample values are typically not preserved through the encoding and decoding processes, although some modes are included that provide lossless coding capability. A number of techniques are specified to enable highly efficient compression. Encoding algorithms (not specified within the scope of this document) may select between inter, intra, intra block copy (IBC), and palette coding for block-shaped regions of each picture. Inter coding uses motion vectors for block-based inter-picture 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 within the same picture, and intra block copy coding uses block displacement vectors to reference previously decoded regions of the same picture to exploit statistical similarities among different areas of the same picture. Motion vectors, intra prediction modes, and IBC block vectors are specified for a variety of block sizes in the picture. The prediction residual can then be further compressed using a spatial transform to remove spatial correlation inside a block before it is quantized, producing a possibly irreversible process that typically discards less important visual information while forming a close approximation to the source samples. Finally, the motion vectors, intra prediction modes, and block vectors can also be further compressed using a variety of prediction mechanisms, and, after prediction, are combined with the quantized transform coefficient information and encoded using arithmetic coding.

## How to read this document

It is suggested that the reader starts with Clause 1 and moves on to Clause 3. Clause 6 should be read for the geometrical relationship of the source, input, and output of the decoder. Clause 7 specifies the order to parse syntax elements from the bitstream. See subclauses 7.1 to 7.3 for syntactical order and subclause 7.4 for semantics; e.g. the scope, restrictions, and conditions that are imposed on the syntax elements. The actual parsing for most syntax elements is specified in Clause 9. Finally, Clause 8 specifies how the syntax elements are mapped into decoded samples. Throughout reading this document, the reader should refer to Clauses 2, 4, and 5 as needed. Annexes A through D also form an integral part of this document.

Annex A specifies profiles, each being tailored to certain application domains, and defines the so-called tiers and 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, bitstream conformance, decoder conformance, and the use of the hypothetical reference decoder to check bitstream and decoder conformance. Annex D specifies syntax and semantics for supplemental enhancement information (SEI) message payloads that affect the conformance specifications in Annex C. Rec. ITU-T H.274 | ISO/IEC 23002-7 specifies the syntax and semantics of the video usability information (VUI) parameters as well as SEI messages that do not affect the conformance specifications in Annex C. These VUI parameters and SEI messages may be used together with this document.

## Patent declarations

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.

ISO and IEC take no position concerning the evidence, validity and scope of these patent rights.

The holders of these patent rights have assured ISO and IEC that they are willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statements of the holders of these patent rights are registered with ISO and IEC. Information may be obtained from the patent database available at www.iso.org/patents.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those in the patent database. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Withdrawn


© ISO/IEC 2021 – All rights reserved

This is a preview - click here to buy the full publication



# Information technology — Coded representation of immersive media — Part 3: Versatile video coding

## 1 Scope

This document specifies a video coding technology known as versatile video coding (VVC), comprising a video coding technology with a compression capability that is substantially beyond that of the prior generations of such standards and with sufficient versatility for effective use in a broad range of applications.

Only the syntax format, semantics, and associated decoding process requirements are specified, while other matters such as pre-processing, the encoding process, system signalling and multiplexing, data loss recovery, post-processing, and video display are considered to be outside the scope of this document. Additionally, the internal processing steps performed within a decoder are also considered to be outside the scope of this document; only the externally observable output behaviour is required to conform to the specifications of this document.

This document is designed to be generic in the sense that it serves a wide range of applications, bit rates, resolutions, qualities and services. Applications include, but are not limited to, video coding for digital storage media, television broadcasting and real-time communication. In the course of creating This document, 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 document is designed to facilitate video data interchange among different applications.

## 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.

Rec. ITU-T H.274 | ISO/IEC 23002-7, *Versatile supplemental enhancement information messages for coded video bitstreams*

Rec. ITU-T T.35:2000, Procedure for the allocation of ITU-T defined codes for non standard facilities.

Withdrawn

© ISO/IEC 2021 – All rights reserved