

Fourth edition
2020-08

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 2: High efficiency video coding

Technologies de l'information — Codage à haute efficacité et livraison des médias dans des environnements hétérogènes —

Partie 2: Codage vidéo à haute efficacité



Reference number
ISO/IEC 23008-2:2020(E)

© ISO/IEC 2020



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2020

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

This is a preview of "ISO/IEC 23008-2:2020". [Click here to purchase the full version from the ANSI store.](#)

Contents	Page
Foreword	vii
Introduction	viii
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	21
5 Conventions	23
5.1 General.....	23
5.2 Arithmetic operators.....	23
5.3 Logical operators.....	23
5.4 Relational operators.....	24
5.5 Bit-wise operators.....	24
5.6 Assignment operators.....	24
5.7 Range notation.....	25
5.8 Mathematical functions.....	25
5.9 Order of operation precedence.....	26
5.10 Variables, syntax elements, and tables.....	27
5.11 Text description of logical operations	28
5.12 Processes.....	29
6 Bitstream and picture formats, partitionings, scanning processes, and neighbouring relationships	30
6.1 Bitstream formats.....	30
6.2 Source, decoded, and output picture formats.....	30
6.3 Partitioning of pictures, slices, slice segments, tiles, CTUs, and CTBs.....	33
6.3.1 Partitioning of pictures into slices, slice segments, and tiles	33
6.3.2 Block and quadtree structures.....	34
6.3.3 Spatial or component-wise partitionings	35
6.4 Availability processes	36
6.4.1 Derivation process for z-scan order block availability.....	36
6.4.2 Derivation process for prediction block availability.....	37
6.5 Scanning processes.....	38
6.5.1 CTB raster and tile scanning conversion process.....	38

6.5.2	Z-scan order array initialization process.....	39
6.5.3	Up-right diagonal scan order array initialization process.....	40
6.5.4	Horizontal scan order array initialization process.....	40
6.5.5	Vertical scan order array initialization process.....	41
6.5.6	Traverse scan order array initialization process.....	41
7	Syntax and semantics.....	41
7.1	Method of specifying syntax in tabular form	41
7.2	Specification of syntax functions and descriptors	43
7.3	Syntax in tabular form.....	45
7.3.1	NAL unit syntax	45
7.3.2	Raw byte sequence payloads, trailing bits, and byte alignment syntax.....	45
7.3.3	Profile, tier and level syntax.....	54
7.3.4	Scaling list data syntax.....	57
7.3.5	Supplemental enhancement information message syntax.....	58
7.3.6	Slice segment header syntax.....	59
7.3.7	Short-term reference picture set syntax.....	63
7.3.8	Slice segment data syntax.....	64
7.4	Semantics.....	80
7.4.1	General	80
7.4.2	NAL unit semantics	80
7.4.3	Raw byte sequence payloads, trailing bits, and byte alignment semantics.....	90
7.4.4	Profile, tier, and level semantics.....	113
7.4.5	Scaling list data semantics	117
7.4.6	Supplemental enhancement information message semantics.....	120
7.4.7	Slice segment header semantics.....	121
7.4.8	Short-term reference picture set semantics.....	130
7.4.9	Slice segment data semantics	133

This is a preview of "ISO/IEC 23008-2:2020". [Click here to purchase the full version from the ANSI store.](#)

8	Decoding process	149
8.1	General decoding process	149
8.1.1	General	149
8.1.2	CVSG decoding process.....	150
8.1.3	Decoding process for a coded picture with nuh_layer_id equal to 0	150
8.2	NAL unit decoding process.....	153
8.3	Slice decoding process.....	153
8.3.1	Decoding process for picture order count.....	153
8.3.2	Decoding process for reference picture set	154
8.3.3	Decoding process for generating unavailable reference pictures	159
8.3.4	Decoding process for reference picture lists construction.....	160
8.3.5	Decoding process for collocated picture and no backward prediction flag	161
8.4	Decoding process for coding units coded in intra prediction mode.....	162
8.4.1	General decoding process for coding units coded in intra prediction mode	162
8.4.2	Derivation process for luma intra prediction mode.....	166
8.4.3	Derivation process for chroma intra prediction mode	169
8.4.4	Decoding process for intra blocks.....	170
8.5	Decoding process for coding units coded in inter prediction mode.....	183
8.5.1	General decoding process for coding units coded in inter prediction mode	183
8.5.2	Inter prediction process.....	184
8.5.3	Decoding process for prediction units in inter prediction mode.....	187
8.5.4	Decoding process for the residual signal of coding units coded in inter prediction mode.....	220
8.6	Scaling, transformation and array construction process prior to deblocking filter process	224
8.6.1	Derivation process for quantization parameters	224
8.6.2	Scaling and transformation process.....	226
8.6.3	Scaling process for transform coefficients	228
8.6.4	Transformation process for scaled transform coefficients	229
8.6.5	Residual modification process for blocks using a transform bypass	232
8.6.6	Residual modification process for transform blocks using cross-component prediction.....	233
8.6.7	Picture construction process prior to in-loop filter process.....	233

8.6.8	Residual modification process for blocks using adaptive colour transform.....	234
8.7	In-loop filter process.....	236
8.7.1	General.....	236
8.7.2	Deblocking filter process.....	237
8.7.3	Sample adaptive offset process.....	255
9	Parsing process	258
9.1	General.....	258
9.2	Parsing process for 0-th order Exp-Golomb codes.....	258
9.2.1	General.....	258
9.2.2	Mapping process for signed Exp-Golomb codes	260
9.3	CABAC parsing process for slice segment data.....	260
9.3.1	General.....	260
9.3.2	Initialization process.....	263
9.3.3	Binarization process.....	277
9.3.4	Decoding process flow.....	287
9.3.5	Arithmetic encoding process	303
10	Sub-bitstream extraction process	309
Annex A (normative)	Profiles, tiers and levels.....	311
Annex B (normative)	Byte stream format.....	339
Annex C (normative)	Hypothetical reference decoder.....	342
Annex D (normative)	Supplemental enhancement information	363
Annex E (normative)	Video usability information	523
Annex F (normative)	Common specifications for multi-layer extensions.....	552
Annex G (normative)	Multiview high efficiency video coding	708
Annex H (normative)	Scalable high efficiency video coding.....	733
Annex I (normative)	3D high efficiency video coding.....	763
Bibliography.....		888

This is a preview of "ISO/IEC 23008-2:2020". Click here to purchase the full version from the ANSI store.

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).

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 the following URL: www.iso.org/iso/foreword.html.

This document was prepared by 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 Rec. ITU-T H.265.

This fourth edition cancels and replaces the third edition (ISO/IEC 23008-2:2017), which has been technically revised. It also incorporates the amendments ISO/IEC 23008-2:2017/Amd 1:2018, ISO/IEC 23008-2:2017/Amd 2:2018, and ISO/IEC 23008-2:2017/Amd 3:2018).

The main changes compared to the previous edition are:

- the specification of two additional profiles (the Monochrome 10 and Main 10 Still Picture profiles);
- the specification of additional supplemental enhancement information (SEI) messages and additional colour-related video usability information codepoint identifiers;
- corrections to various minor defects.

A list of all parts in the ISO/IEC 23008 series can be found on the ISO website.

Introduction

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 Collaborative Team on Video Coding (JCT-VC) in 2010 and a Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) in 2012 for development of a new Recommendation | International Standard. This document was developed in the JCT-VC and the JCT-3V.

Purpose

This document 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 communications. It is also designed to enable the use of the coded video representation in a flexible manner for a wide variety of network environments as well as to enable the use of multi-core parallel encoding and decoding devices. The use of this document 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. Supports for higher bit depths and enhanced chroma formats, including the use of full-resolution chroma are provided. Support for scalability enables video transmission on networks with varying transmission conditions and other scenarios involving multiple bit rate services. Support for multiview enables representation of video content with multiple camera views and optional auxiliary information. Support for 3D enables joint representation of video content and depth information with multiple camera views.

Applications

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

- broadcast (cable TV on optical networks / copper, satellite, terrestrial, etc.);
- camcorders;
- content production and distribution;
- digital cinema;
- home cinema;
- internet streaming, download and play;
- medical imaging;
- mobile streaming, broadcast and communications;
- real-time conversational services (videoconferencing, videophone, telepresence, etc.);
- remote video surveillance;
- storage media (optical disks, digital video tape recorder, etc.);
- wireless display.

This is a preview of "ISO/IEC 23008-2:2020". [Click here to purchase the full version from the ANSI store.](#)

Publication and versions of this document

This document has been jointly developed by ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG). It is published as technically-aligned twin text in both ITU-T and ISO/IEC. As the basis text has been drafted to become both an ITU-T Recommendation and an ISO/IEC International Standard, the term "Specification" (with capitalization to indicate that it refers to the whole of the text) is used herein when the text refers to itself.

This is the fifth version of this document and the fourth edition published by ISO/IEC.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 1 refers to the first approved version of this document. The first edition published by ISO/IEC as ISO/IEC 23008-2:2013 corresponded to the first version.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 2 refers to the integrated text additionally containing format range extensions, scalability extensions, multiview extensions, additional supplement enhancement information, and corrections to various minor defects in the prior content of the specification. The second edition published by ISO/IEC as ISO/IEC 23008-2:2015 corresponded to the second version.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 3 refers to the integrated text additionally containing 3D extensions, additional supplement enhancement information, and corrections to various minor defects in the prior content of the specification.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 4 refers to the integrated text additionally containing screen content coding extensions profiles, scalable range extensions profiles, additional high throughput profiles, additional supplement enhancement information, additional colour representation identifiers, and corrections to various minor defects in the prior content of the specification. The third edition published by ISO/IEC as ISO/IEC 23008-2:2017 corresponded to the fourth version.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 5 refers to the integrated text additionally containing additional SEI messages that include omnidirectional video specific SEI messages, a Monochrome 10 profile, a Main 10 Still Picture profile, and corrections to various minor defects in the prior content of the specification.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 6 refers to the integrated text additionally containing additional SEI messages for SEI manifest and SEI prefix, and corrections to various minor defects in the prior content of the specification.

Rec. ITU-T H.265 | ISO/IEC 23008-2 version 7 (the current version) refers to the integrated text additionally containing the fisheye video information SEI message and the annotated regions SEI message, and corrections to various minor defects in the prior content of the specification. The fourth edition published by ISO/IEC as ISO/IEC 23008-2:2020 corresponds to the fifth version.

Profiles, tiers, and levels

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 should cover, among other things, digital storage media, television broadcasting, and real-time communications. 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 will 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. 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). 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.

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 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 Recommendation | 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 may then be further compressed using a transform to remove spatial correlation inside the transform 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 or intra prediction modes may 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 (Scope) and moves on to Clause 3 (Terms and 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 7.1 to 7.3 for syntactical order and see 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 (Parsing process). Clause 10 (Sub-bitstream extraction process) specifies the sub-bitstream extraction process. Finally, Clause 8 (Decoding process) specifies how the syntax elements are mapped into decoded samples. Throughout reading this document, the reader should refer to Clauses 2 (Normative references), 4 (Abbreviations), and 5 (Conventions) as needed. Annexes A through I 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

This is a preview of "ISO/IEC 23008-2:2020". [Click here to purchase the full version from the ANSI store.](#)

enhancement information message payloads. Annex E specifies syntax and semantics of the video usability information parameters of the sequence parameter set. Annex F specifies general multi-layer support for bitstreams and decoders. Annex G contains support for multiview coding. Annex H contains support for scalability. Annex I contains support for 3D coding.

Throughout this document, statements appearing with the preamble "NOTE" are informative and are not an integral part of this document.

Patent rights identified

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.