

---

---

**Information technology — MPEG video  
technologies —**

**Part 7:  
Versatile supplemental enhancement  
information messages for coded video  
bitstreams**

*Technologies de l'information — Technologies vidéo MPEG —*

*Partie 7: Messages d'améliorations complémentaires polyvalents pour  
les flux binaires vidéo codés*

ISO/IEC 23002-7:2021

<https://standards.iteh.ai/catalog/standards/iso/4ad46857-bb5d-4b88-ba8a-a336bf8c9e20/iso-iec-23002-7-2021>



iTeh Standards  
(<https://standards.iteh.ai>)  
Document Preview

ISO/IEC 23002-7:2021

<https://standards.iteh.ai/catalog/standards/iso/4ad46857-bb5d-4b88-ba8a-a336bf8c9e20/iso-iec-23002-7-2021>



**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](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

<b>Foreword</b>	<b>v</b>
<b>Introduction</b>	<b>vi</b>
<b>1 Scope</b>	<b>1</b>
<b>2 Normative references</b>	<b>1</b>
<b>3 Terms and definitions</b>	<b>1</b>
<b>4 Abbreviated terms</b>	<b>8</b>
<b>5 Conventions</b>	<b>9</b>
5.1 General	9
5.2 Arithmetic operators	9
5.3 Logical operators	10
5.4 Relational operators	10
5.5 Bit-wise operators	10
5.6 Assignment operators	11
5.7 Range notation	11
5.8 Mathematical functions	11
5.9 Order of operation precedence	12
5.10 Variables, syntax elements and tables	13
5.11 Text description of logical operations	14
5.12 Processes	15
<b>6 Syntax and semantics</b>	<b>16</b>
6.1 General	16
6.2 Method of specifying syntax in tabular form	17
6.3 Specification of syntax functions and descriptors	18
<b>7 Video usability information parameters</b>	<b>18</b>
7.1 General	18
7.2 VUI parameters syntax	19
7.3 VUI parameters semantics	19
<b>8 SEI messages</b>	<b>27</b>
8.1 General	27
8.2 Filler payload SEI message	28
8.2.1 Filler payload SEI message syntax	28
8.2.2 Filler payload SEI message semantics	28
8.3 User data registered by Recommendation ITU-T T.35 SEI message	28
8.3.1 User data registered by Recommendation ITU-T T.35 SEI message syntax	28
8.3.2 User data registered by Recommendation ITU-T T.35 SEI message semantics	28
8.4 User data unregistered SEI message	29
8.4.1 User data unregistered SEI message syntax	29
8.4.2 User data unregistered SEI message semantics	29
8.5 Film grain characteristics SEI message	29
8.5.1 Film grain characteristics SEI message syntax	29
8.5.2 Film grain characteristics SEI message semantics	30
8.6 Frame packing arrangement SEI message	38
8.6.1 Frame packing arrangement SEI message syntax	38
8.6.2 Frame packing arrangement SEI message semantics	38
8.7 Parameter sets inclusion indication SEI message	47
8.7.1 Parameter sets inclusion indication SEI message syntax	47
8.7.2 Parameter sets inclusion indication SEI message semantics	47
8.8 Decoded picture hash SEI message	48
8.8.1 Decoded picture hash SEI message syntax	48
8.8.2 Decoded picture hash SEI message semantics	48
8.9 Mastering display colour volume SEI message	50

8.9.1	Mastering display colour volume SEI message syntax.....	50
8.9.2	Mastering display colour volume SEI message semantics.....	50
8.10	Content light level information SEI message.....	52
8.10.1	Content light level information SEI message syntax.....	52
8.10.2	Content light level information SEI message semantics.....	52
8.11	Dependent random access point indication SEI message.....	53
8.11.1	Dependent random access point indication SEI message syntax.....	53
8.11.2	Dependent random access point indication SEI message semantics.....	53
8.12	Alternative transfer characteristics information SEI message.....	54
8.12.1	Alternative transfer characteristics information SEI message syntax.....	54
8.12.2	Alternative transfer characteristics SEI message semantics.....	54
8.13	Ambient viewing environment SEI message.....	54
8.13.1	Ambient viewing environment SEI message syntax.....	54
8.13.2	Ambient viewing environment SEI message semantics.....	54
8.14	Content colour volume SEI message.....	55
8.14.1	Content colour volume SEI message syntax.....	55
8.14.2	Content colour volume SEI message semantics.....	55
8.15	Omnidirectional video specific SEI messages.....	58
8.15.1	Sample location remapping process.....	58
8.15.2	Equirectangular projection SEI message.....	68
8.15.3	Generalized cubemap projection SEI message.....	70
8.15.4	Sphere rotation SEI message.....	76
8.15.5	Region-wise packing SEI message.....	78
8.15.6	Omnidirectional viewport SEI message.....	85
8.16	Frame-field information SEI message.....	87
8.16.1	Frame-field information SEI message syntax.....	87
8.16.2	Frame-field information SEI message semantics.....	87
8.17	Sample aspect ratio information SEI message.....	90
8.17.1	Sample aspect ratio information SEI message syntax.....	90
8.17.2	Sample aspect ratio information SEI message semantics.....	90
8.18	Reserved SEI message.....	91
8.18.1	Reserved SEI message syntax.....	91
8.18.2	Reserved SEI message semantics.....	91
<b>9</b>	<b>Parsing process for k-th order Exp-Golomb codes.....</b>	<b>91</b>
9.1	General.....	91
9.2	Mapping process for signed Exp-Golomb codes.....	92
	<b>Bibliography.....</b>	<b>94</b>

## 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](http://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](http://www.iso.org/patents)) or the IEC list of patent declarations received (see [patents.iec.ch](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](http://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 Recommendation ITU-T H.274 (08/2020).

A list of all parts in the ISO/IEC 23002 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](http://www.iso.org/members.html).

## Introduction

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](http://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.

iTeh Standards  
(<https://standards.iteh.ai>)  
Document Preview

ISO/IEC 23002-7:2021

<https://standards.iteh.ai/catalog/standards/iso/4ad46857-bb5d-4b88-ba8a-a336bf8c9e20/iso-iec-23002-7-2021>

# Information technology — MPEG video technologies —

## Part 7:

# Versatile supplemental enhancement information messages for coded video bitstreams

## 1 Scope

This document specifies the syntax and semantics of video usability information (VUI) parameters and supplemental enhancement information (SEI) messages. The VUI parameters and SEI messages defined in this document are designed to be conveyed within coded video bitstreams in a manner specified in a video coding specification or to be conveyed by other means determined by the specifications for systems that make use of such coded video bitstreams. This document is particularly intended for use with coded video bitstreams as specified by Rec. ITU-T H.266 | ISO/IEC 23090-3, although it can also be used with other types of coded video bitstreams.

VUI parameters and SEI messages can assist in processes related to decoding, display or other purposes. However, unless otherwise specified in a referencing specification, the interpretation and use of the VUI parameters and SEI messages specified in this document is not a required functionality of a video decoder or receiving video system. Although semantics are specified for the VUI parameters and SEI messages, decoders and receiving video systems can simply ignore the content of the VUI parameters and SEI messages or can use them in a manner that somewhat differs from what is specified in this document.

## 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.

Recommendation ITU-T H.273 | ISO/IEC 23091-2, *Coding-independent code points for video signal type identification*

ISO/CIE 11664-1, *Colorimetry — Part 1: CIE standard colorimetric observers*

ISO/IEC 11578:1996, *Information technology — Open Systems Interconnection — Remote Procedure Call (RPC)*

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

IETF RFC 1321, *The MD5 Message-Digest Algorithm*

## 3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

### 3.1

#### AU

access unit

set of *PUs* that belong to different *layers* and contain *coded pictures* associated with the same *output time*

### 3.2

#### APS

adaptation parameter set

*syntax structure* containing *syntax elements* that apply to zero or more *slices* as determined by zero or more *syntax elements* found in *slice headers*

### 3.3

#### associated IRAP picture

previous *IRAP picture* (when present) in *decoding order*, for a particular picture, in the same *layer* as the particular *picture*

### 3.4

#### azimuth circle

circle on a sphere connecting all points with the same azimuth value

Note 1 to entry: An azimuth circle is always a *great circle* like a longitude line on the earth.

### 3.5

#### byte

sequence of 8 bits, within which, when written or read as a sequence of bit values, the left-most and right-most bits represent the most and least significant bits, respectively, and the bits are written or read from left to right

### 3.6

#### chroma

sample array or single sample representing one of the two colour difference signals related to the primary colours, represented by the symbols Cb and Cr

Note 1 to entry: The term chroma is used rather than the term chrominance in order to avoid implying the use of linear light transfer characteristics that is often associated with the term chrominance.

### 3.7

#### CLVS

coded layer video sequence

sequence of *PUs* of the same layer that consists, in *decoding order*, of a *CLVSS PU*, followed by zero or more *PUs* that are not *CLVSS PUs*, including all subsequent *PUs* up to but not including any subsequent *PU* that is a *CLVSS PU*

### 3.8

#### CLVSS PU

*PU* in which the *coded picture* is a *CLVSS picture*

### 3.9

#### CLVSS

coded layer video sequence start

*coded picture* that starts a new CLVS as specified in a video coding specification

Note 1 to entry: In Rec. ITU-T H.266 | ISO/IEC 23090-3, a CLVSS picture is an *IRAP picture* with NoIncorrectPicOutputFlag equal to 1 or a gradual decoding refresh picture with NoIncorrectPicOutputFlag equal to 1. In Rec. ITU-T H.265 | ISO/IEC 23008-2, a CLVSS picture is an *IRAP picture* with NoRaslOutputFlag equal to 1.

### 3.10

#### coded picture

*coded representation* of a *picture* containing all *CTUs* of the *picture*



**3.11****coded slice NAL unit**

*NAL unit* that contains a coded *slice*

**3.12****coded video bitstream**

sequence of bits that forms the representation of a sequence of *AUs* forming one or more *CVSs*

**3.13****CVS****coded video sequence**

sequence of *AUs* that consists, in *decoding order*, of a *CVSS AU*, followed by zero or more *AUs* that are not *CVSS AUs*, including all subsequent *AUs* up to but not including any subsequent *AU* that is a *CVSS AU*

**3.14****CVSS AU**

*AU* that has a *PU* for each *layer* present in the *CVS* and the coded *picture* in each *PU* is a *CLVSS picture*

**3.15****component**

array or single sample from one of the three arrays (*luma* and two *chroma*) that compose a *picture* in 4:2:0, 4:2:2, or 4:4:4 colour format or the array or a single sample of the array that compose a *picture* in monochrome format

**3.16****constituent picture**

part of a spatially *frame-packed* stereoscopic *picture* that corresponds to one view, or a *picture* itself when *frame packing* is not in use or the temporal interleaving *frame packing* arrangement is in use

**3.17****cropped decoded picture**

result of cropping a *decoded picture* based on the conformance cropping window for the corresponding *coded picture*

**3.18****decoded picture**

*decoded picture* is derived by decoding a *coded picture*

**3.19****decoder**

embodiment of a *decoding process*

**3.20****decoding order**

order in which *syntax elements* are conveyed in the *coded video bitstream* and are processed by a *decoding process*

**3.21****decoding process**

process that reads a *coded video bitstream* and derives *decoded pictures* from it

**3.22****elevation circle**

circle on a sphere connecting all points with the same elevation value

Note 1 to entry: An elevation circle is similar to a latitude line on the earth. Except when the elevation value is zero, an elevation circle is not a *great circle* like a longitude circle on the earth.

**3.23****encoder**

embodiment of an *encoding process*

**3.24**

**encoding process**

process that produces a *coded video bitstream*

**3.25**

**field**

assembly of alternative rows of samples of a *frame*

**3.26**

**flag**

variable or single-bit *syntax element* that can take one of the two possible values: 0 and 1

**3.27**

**frame**

composition of a top *field* and a bottom *field*, where sample rows 0, 2, 4, ... originate from the top *field* and sample rows 1, 3, 5, ... originate from the bottom *field*

**3.28**

**global coordinate axes**

coordinate axes associated with *omnidirectional video* that are associated with an externally referenceable position and orientation

Note 1 to entry: The global coordinate axes could correspond to the position and orientation of a device or rig used for omnidirectional audio/video acquisition as well as the position of an observer's head in the three-dimensional space of the *omnidirectional video* rendering environment.

**3.29**

**great circle**

intersection of a sphere and a plane that passes through the centre point of the sphere

Note 1 to entry: A great circle is also known as an orthodrome or Riemannian circle.

**3.30**

**inter prediction**

aspect of the *decoding process* for a *coded picture* that makes use of data derived from the *decoding process* of one or more previously decoded *reference pictures*

**3.31**

**IRAP picture**

*coded picture* starting from which all *pictures* in the same *layer* in both *decoding order* and *output order* can be decoded without first decoding any *picture* in the same *layer* earlier in *decoding order* in the *coded video bitstream*

**3.32**

**layer**

set of *VCL NAL units* that all have a particular value of layer identifier and the associated non-VCL *NAL units*, wherein the layer identifier is a variable for which the value is specified by a video coding specification

Note 1 to entry: In the contexts of Rec. ITU-T H.266 | ISO/IEC 23090-3 and Rec. ITU-T H.265 | ISO/IEC 23008-2, the layer identifier is the value of the *nuh\_layer\_id* syntax element in the *NAL unit header*.

**3.33**

**leading picture**

*picture* that is in the same *layer* as the *associated IRAP picture* and precedes the *associated IRAP picture* in *output order*

**3.34**

**local coordinate axes**

coordinate axes having a specified rotation relationship relative to the *global coordinate axes*

**3.35****luma**

sample array or single sample representing the monochrome signal related to the primary colours, represented by the symbol or subscript Y or L

Note 1 to entry: The term luma is used rather than the term luminance in order to avoid implying the use of linear light transfer characteristics that is often associated with the term luminance. The symbol L is sometimes used instead of the symbol Y to avoid confusion with the symbol y as used for vertical location.

**3.36****NAL unit**

*syntax structure* containing an indication of the type of data that follows and *bytes* containing that data in a manner that enables the extraction of a string of data bits from the syntax structure

**3.37****non-VCL NAL unit**

*NAL unit* that is not a *VCL NAL unit*

**3.38****omnidirectional video**

video content in a format that enables rendering according to the user's viewing orientation, e.g., if viewed using a head-mounted device, or according to a user's desired *viewport*, reflecting a potentially rotated viewing position

**3.39****output order**

order in which the *decoded pictures* are output from the *decoder* (for the *decoded pictures* that are to be output from the *decoder*)

**3.40****output time**

time when a *decoded picture* is to be output from the *decoder* (for the *decoded pictures* that are to be output from the *decoder*)

**3.41****packed region**

region in a *region-wise packed picture* that is mapped to a *projected region* according to a *region-wise packing*

**3.42****picture**

array of *luma* samples in monochrome format or an array of *luma* samples and two corresponding arrays of *chroma* samples in 4:2:0, 4:2:2, and 4:4:4 colour format

Note 1 to entry: A picture could be either a frame or a field. However, in one CLVS, either all pictures are frames or all pictures are fields.

**3.43****PPS****picture parameter set**

A *syntax structure* containing *syntax elements* that apply to zero or more entire *coded pictures* as determined by a *syntax element* that is the same for all *slices* of a picture and found in the picture header or *slice headers* of each *picture*

**3.44****PU****picture unit**

set of *NAL units* that contain all *VCL NAL units* of a *coded picture* and their associated non-VCL NAL units

**3.45**

**projected picture**

*picture that uses a projection format for omnidirectional video*

**3.46**

**projected region**

*region in a projected picture that is mapped to a packed region according to a region-wise packing*

**3.47**

**projection**

*specified correspondence between the colour samples of a projected picture and azimuth and elevation positions on a sphere*

**3.48**

**random access**

*act of starting the decoding process for a coded video bitstream at a point other than the beginning of the bitstream*

**3.49**

**RASL picture**

*leading picture that cannot be correctly decoded when the decoding process starts from the associated IRAP picture*

**3.50**

**reference picture**

*picture that contains samples that could be used for inter prediction in the decoding process of subsequent pictures in decoding order*

**3.51**

**reference picture list**

*list of reference pictures that is used for inter prediction of a slice*

**3.52**

**region-wise packed picture**

*decoded picture that contains one or more packed regions*

Note 1 to entry: A region-wise packed picture could contain a *region-wise packing* of a *projected picture*.

**3.53**

**region-wise packing**

*transformation, resizing, and relocation of packed regions of a region-wise packed picture to remap the packed regions to projected regions of a projected picture*

**3.54**

**sample aspect ratio**

*indicated width-to-height aspect ratio of the luma samples of the associated decoded pictures*

**3.55**

**slice**

*region of a picture that can be decoded separately from other regions of the same coded picture (although in some cases the decoding process for the picture might use inter prediction that makes reference to other previously decoded reference pictures)*

**3.56**

**source**

*term used to describe the video material or some of its attributes before encoding*

**3.57**

**sphere coordinates**

*azimuth and elevation angles identifying a location of a point on a sphere*

**3.58****sphere region**

region on a sphere, specified either by four *great circles* or by two *azimuth circles* and two *elevation circles*, or such a region on a rotated sphere after applying yaw, pitch, and roll rotations

**3.59****STSA picture**

*coded picture* that enables up-switching, at the *coded picture*, to the *temporal sublayer* containing the *coded picture*, from the immediately lower *temporal sublayer* of the same layer when the *coded picture* does not belong to the lowest *temporal sublayer*

Note 1 to entry: An STSA picture does not use pictures in the same layer and with the same temporal sublayer identifier as the STSA picture for inter prediction reference. Pictures following an STSA picture in decoding order in the same layer and with the same temporal sublayer identifier as the STSA picture do not use pictures prior to the STSA picture in decoding order in the same layer and with the same temporal sublayer identifier as the STSA picture for inter prediction reference. STSA pictures in an independent layer (i.e., a layer that does not depend on other layers in its decoding) always have a temporal sublayer identifier greater than 0.

**3.60****syntax element**

element of data represented in a *syntax structure*

**3.61****syntax structure**

zero or more *syntax elements* that are present together in a specified order in a string of data bits, where the left-most bit is considered to be the first and most significant bit, and the right-most bit is considered to be the last and least significant bit

**3.62****SEI message**

*syntax structure* that provides a particular type of information that assists in processes related to decoding, display or other purposes but is not needed by the *decoding process* in order to determine the values of the samples in *decoded pictures*

**3.63****temporal sublayer**

subset of a temporal scalable *bitstream*, consisting of *VCL NAL units* with a particular value of *temporal sublayer identifier* and the associated *non-VCL NAL units*

**3.64****temporal sublayer identifier**

number greater than or equal to 0 defined by a variable for which the value is specified by a video coding specification such that pictures of all *temporal sublayers* have a specified temporal output order relative to each other and pictures with a lower temporal sublayer identifier can be decoded without reference to pictures with a higher temporal sublayer identifier

**3.65****tilt angle**

angle indicating the amount of tilt of a *sphere region*, measured as the amount of rotation of a *sphere region* along the axis originating from the sphere origin passing through the centre point of the *sphere region*, where the angle value increases clockwise when looking from the origin towards the positive end of the axis

**3.66****trailing picture**

*coded picture* that is associated with an *IRAP picture* and follows the *IRAP picture* in both decoding order and output order

**3.67**

**VCL NAL unit**

collective term for *coded slice NAL units* and the subset of other *NAL units* that have *reserved* values of *NAL unit* type identifiers that are classified as VCL NAL units in a referencing specification

**3.68**

**VUI parameters**

syntax structure that identifies properties of interpretation of decoded pictures for display purposes, particularly including colour representation information

**3.69**

**viewport**

region of *omnidirectional video* content suitable for display and viewing by the user

## **4 Abbreviated terms**

AU	access unit
APS	adaptation parameter set
CRC	cyclic redundancy check
CLVS	coded layer video sequence
CLVSS	coded layer video sequence start
CVS	coded video sequence
DRAP	dependent random access point
FIR	finite impulse response
IRAP	intra random access point
NAL	network abstraction layer
PPS	picture parameter set
PU	picture unit
RASL	random access skipped leading
RWP	region-wise packing
SAR	sample aspect ratio
SARI	sample aspect ratio information
SEI	supplemental enhancement information
STSA	step-wise temporal sublayer access
VCL	video coding layer
VUI	video usability information