
**Information technology — Coded
representation of immersive media —
Part 5:
Visual volumetric video-based coding
(V3C) and video-based point cloud
compression (V-PCC)**

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

*Partie 5: Codage basé sur la vidéo volumétrique (V3C) et compression
de nuage de points basée sur la vidéo (V-PCC)*





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

Contents

	Page
Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	2
4 Abbreviated terms	15
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.....	18
5.7 Other operators.....	18
5.8 Mathematical functions.....	18
5.9 Order of operation precedence.....	19
5.10 Variables, syntax elements and tables.....	19
5.11 Text description of logical operations.....	21
5.12 Processes.....	22
6 Overall V3C characteristics, decoding operations and post-decoding processes	23
6.1 V3C characteristics.....	23
6.2 V3C bitstream characteristics, decoding operations and post-decoding processes.....	25
7 Bitstream format, partitioning and scanning processes	26
7.1 General.....	26
7.2 V3C bitstream formats.....	26
7.3 NAL bitstream formats.....	27
7.4 Partitioning of atlas frames into tiles.....	27
7.5 Tile partition scanning process.....	28
8 Syntax and semantics	29
8.1 Method of specifying syntax in tabular form.....	29
8.2 Specification of syntax functions and descriptors.....	30
8.3 Syntax in tabular form.....	32
8.3.1 General.....	32
8.3.2 V3C unit syntax.....	33
8.3.3 Byte alignment syntax.....	34
8.3.4 V3C parameter set syntax.....	35
8.3.5 NAL unit syntax.....	38
8.3.6 Raw byte sequence payloads, trailing bits and byte alignment syntax.....	38
8.3.7 Atlas tile data unit syntax.....	44
8.3.8 Supplemental enhancement information message syntax.....	48
8.4 Semantics.....	49
8.4.1 General.....	49
8.4.2 V3C unit semantics.....	49
8.4.3 Byte alignment semantics.....	51
8.4.4 V3C parameter set semantics.....	51
8.4.5 NAL unit semantics.....	57
8.4.6 Raw byte sequence payloads, trailing bits and byte alignment semantics.....	66
8.4.7 Atlas tile data unit semantics.....	79
8.4.8 Supplemental enhancement information message semantics.....	87
9 Decoding process	87
9.1 General decoding process.....	87

9.2	Atlas data decoding process	88
9.2.1	General atlas data decoding process	88
9.2.2	Decoding process for a coded atlas frame	89
9.2.3	Atlas NAL unit decoding process	90
9.2.4	Atlas tile header decoding process	90
9.2.5	Decoding process for patch data units	95
9.2.6	Decoding process of the block to patch map	110
9.2.7	Conversion of tile level patch information to atlas level patch information	111
9.3	Occupancy video decoding process	113
9.4	Geometry video decoding process	114
9.5	Attribute video decoding process	116
9.6	Sub-bitstream extraction process	119
9.6.1	General	119
9.6.2	V3C unit extraction	119
9.6.3	NAL unit extraction process	120
10	Pre-reconstruction process	120
11	Reconstruction process	120
12	Post-reconstruction process	120
13	Adaptation process	120
14	Parsing process	121
14.1	General	121
14.2	Parsing process for 0-th order Exp-Golomb codes	121
14.2.1	General	121
14.2.2	Mapping process for signed Exp-Golomb codes	122
Annex A	(normative) Profiles, tiers and levels	123
Annex B	(normative) Post-decoding conversion to nominal video formats	137
Annex C	(normative) V3C sample stream format	158
Annex D	(normative) NAL sample stream format	160
Annex E	(normative) Atlas hypothetical reference decoder	162
Annex F	(normative) Supplemental enhancement information	179
Annex G	(normative) Volumetric usability information	236
Annex H	(normative) Video-based point cloud coding	247
Bibliography	331

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*.

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.

Introduction

Advances in 3D capturing and rendering technologies have unleashed a new wave of innovation in virtual/augmented/mixed reality (VR/AR/MR) content creation and communication, of which visual volumetric video is an integral part.

Visual volumetric video, a sequence of visual volumetric frames, if uncompressed, can be represented by a large amount of data, which can be costly in terms of storage and transmission. This has led to the need for a high coding efficiency standard for the compression of visual volumetric data.

Visual volumetric frames can be coded by converting the 3D volumetric information into a collection of 2D images and associated data. The converted 2D images can be coded using widely available video and image coding specifications, such as ISO/IEC 14496-10 and ISO/IEC 23008-2 and the associated data can be coded with mechanisms specified in this document. The coded images and the associated data can then be decoded and used to reconstruct the 3D volumetric information. This document specifies a generic mechanism for visual volumetric video coding, i.e. visual volumetric video-based coding. The generic mechanism can be used by applications targeting volumetric content, such as point clouds, immersive video with depth, mesh representations of visual volumetric frames, etc.

In addition to the generic mechanism of coding volumetric content, this document specifies one of the applications of visual volumetric video-based coding targeting point cloud representations of visual volumetric frames. In a point cloud sequence, each point cloud frame contains a collection of points. Each point has a 3D position, i.e., geometry information, and each point can also be associated with a number of attributes, such as colour, reflectance, surface normal, etc.

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 a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is 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.

Information technology — Coded representation of immersive media —

Part 5:

Visual volumetric video-based coding (V3C) and video-based point cloud compression (V-PCC)

1 Scope

This document specifies the syntax, semantics and decoding for visual volumetric media using video-based coding methods. This document also specifies processes that can be needed for reconstruction of visual volumetric media, which can also include additional processes such as post-decoding, pre-reconstruction, post-reconstruction and adaptation.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitute 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.

IEEE 754-2019, *IEEE Standard for Floating-Point Arithmetic*

IETF RFC 1321 *The MD5 Message-Digest Algorithm*

IETF RFC 5646 *Tags for Identifying Languages*

ISO/IEC 10646, *Information technology — Universal Coded Character Set (UCS)*

ISO/IEC 14496-10:2020, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 14496-15, *Information technology — Coding of audio-visual objects — Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format*

ISO/IEC 23008-2:2020, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

ISO/IEC 23090-3, *Information technology — Coded representation of immersive media — Part 3: Versatile video coding*

ISO/IEC 23091-2, *Coding-independent code points for video signal type identification*

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

Rec. ITU-T H.271:2006, *Video back-channel messages for conveyance of status information and requests from a video receiver to a video sender*