A picture containing drawing, food

Description automatically generated

**ISO/IEC JTC 1/SC 29/WG 11 N19089**

**ISO/IEC JTC 1/SC 29/WG 11**

**Coding of moving pictures and audio**

**Convenorship: UNI (Italy)**

**Document type: Approved WG 11 document**

**Title: V-PCC Future Enhancements**

**Status: Approved**

**Date of document: 2020-04-14**

**Source: 3DG**

**Expected action:**

**No. of pages: 221**

**Email of convenor: leonardo@chiariglione.org**

**Committee URL: mpeg.chiariglione.org**

**ISO/IEC 23090-5:2019(E)**



ISO/IEC JTC 1/SC 29/WG 11

Secretariat: JISC

**Information technology — Coded Representation of Immersive Media — Part 5: Video-based Point Cloud Compression**

DIS stage

**Warning for WDs and CDs**

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

© ISO/IEC 2019

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

Fax: +41 22 749 09 47

Email: copyright@iso.org

Website: www.iso.org

Published in Switzerland

Contents

[Foreword viii](#_Toc35407557)

[Introduction ix](#_Toc35407558)

[1 Scope 1](#_Toc35407559)

[2 Normative reference 1](#_Toc35407560)

[3 Terms and definitions 1](#_Toc35407561)

[4 Abbreviations 8](#_Toc35407562)

[5 Conventions 9](#_Toc35407563)

[5.1 General 9](#_Toc35407564)

[5.2 Arithmetic operators 9](#_Toc35407565)

[5.3 Logical operators 10](#_Toc35407566)

[5.4 Relational operators 10](#_Toc35407567)

[5.5 Bit-wise operators 10](#_Toc35407568)

[5.6 Assignment operators 11](#_Toc35407569)

[5.7 Range notation 11](#_Toc35407570)

[5.8 Mathematical functions 11](#_Toc35407571)

[5.9 Order of operation precedence 12](#_Toc35407572)

[5.10 Variables, syntax elements, and tables 13](#_Toc35407573)

[5.11 Text description of logical operations 14](#_Toc35407574)

[5.12 Processes 15](#_Toc35407575)

[6 Bitstream format, partitioning, and scanning processes 15](#_Toc35407576)

[6.1 V-PCC bitstream formats 15](#_Toc35407577)

[6.2 NAL bitstream formats 16](#_Toc35407578)

[6.3 Partitioning of atlas frames, tiles, and tile groups 16](#_Toc35407579)

[6.3.1 Partitioning of atlas frames into tiles and tile groups 16](#_Toc35407580)

[6.4 Scanning Processes 17](#_Toc35407581)

[6.4.1 Tile scanning process 17](#_Toc35407582)

[7 Syntax and semantics 18](#_Toc35407583)

[7.1 Method of specifying syntax in tabular form 18](#_Toc35407584)

[7.2 Specification of syntax functions and descriptors 19](#_Toc35407585)

[7.3 Syntax in tabular form 22](#_Toc35407586)

[7.3.1 General 22](#_Toc35407587)

[7.3.2 V-PCC unit syntax 23](#_Toc35407588)

[7.3.3 Byte alignment syntax 24](#_Toc35407589)

[7.3.4 Sequence parameter set syntax 24](#_Toc35407590)

[7.3.5 NAL unit syntax 26](#_Toc35407591)

[7.3.6 Atlas sequence, frame, and tile group parameter set syntax 26](#_Toc35407592)

[7.3.7 Atlas tile group data unit syntax 32](#_Toc35407593)

[7.3.8 Supplemental enhancement information message syntax 36](#_Toc35407594)

[7.4 Semantics 37](#_Toc35407595)

[7.4.1 General 37](#_Toc35407596)

[7.4.2 V-PCC unit semantics 37](#_Toc35407597)

[7.4.3 Byte alignment semantics 38](#_Toc35407598)

[7.4.4 V-PCC parameter set semantics 38](#_Toc35407599)

[7.4.5 NAL unit semantics 43](#_Toc35407600)

[7.4.6 Atlas sequence, frame, and tile group parameter set semantics 50](#_Toc35407601)

[7.4.7 Atlas tile group data unit semantics 65](#_Toc35407602)

[7.4.8 Supplemental enhancement information message semantics 73](#_Toc35407603)

[8 Decoding process 74](#_Toc35407604)

[8.1 General decoding process 74](#_Toc35407605)

[8.2 Occupancy video decoding process 75](#_Toc35407606)

[8.3 Geometry video decoding process 75](#_Toc35407608)

[8.4 Attribute video decoding process 76](#_Toc35407613)

[8.5 Atlas data decoding process 76](#_Toc35407619)

[8.5.1 General atlas data decoding process 76](#_Toc35407620)

[8.5.2 Atlas data unit decoding process 77](#_Toc35407621)

[8.5.3 Atlas data header decoding process 77](#_Toc35407622)

[8.5.4 Decoding process for patch units 80](#_Toc35407623)

[8.5.5 Decoding process of the block to patch map 89](#_Toc35407624)

[8.5.6 Conversion of tile group level information to Atlas level information 90](#_Toc35407625)

[8.6 Occupancy map video decoding process 90](#_Toc35407627)

[9 Reconstruction process 91](#_Toc35407628)

[9.1 General 91](#_Toc35407629)

[9.2 Chroma format, resolution, and frame rate conversion 92](#_Toc35407630)

[9.3 Point cloud reconstruction 93](#_Toc35407631)

[9.4 Reconstruction of points for non-RAW intra and inter coded patches 95](#_Toc35407632)

[9.4.1 Reconstruction of points for non-RAW intra and inter coded patches when asps\_pixel\_deinterleaving\_enabled\_flag is equal to 1 95](#_Toc35407633)

[9.4.2 Reconstruction of points for non-RAW intra and inter coded patches when plri\_point\_local\_reconstruction\_map\_flag is equal to 1 98](#_Toc35407634)

[9.4.3 Reconstruction of points for non-RAW intra and inter coded patches 102](#_Toc35407635)

[9.4.4 Reconstruction of points when asps\_eom\_patch\_enabled\_flag is 1 103](#_Toc35407636)

[9.4.5 Conversion from patch coordinates to atlas coordinates 105](#_Toc35407637)

[9.4.6 Reconstruct 3D point position 106](#_Toc35407638)

[9.4.7 Projection planes at 45 degrees 106](#_Toc35407639)

[9.4.8 Duplicate point check 107](#_Toc35407640)

[9.5 Reconstruction of point cloud from RAW coded patches 107](#_Toc35407641)

[9.6 Smoothing process 109](#_Toc35407642)

[9.6.1 General 109](#_Toc35407643)

[9.6.2 Geometry smoothing process 110](#_Toc35407644)

[9.6.3 Attribute smoothing process 112](#_Toc35407684)

[9.6.4 Identification of the Center Grid 115](#_Toc35407685)

[9.6.5 Identification of Boundary Points 119](#_Toc35407687)

[9.6.6 Identification of 2 × 2 ×2 cell for a reconstructed point position 119](#_Toc35407690)

[9.6.7 Trilinear Filtering 120](#_Toc35407691)

[9.6.8 Patch Border Filtering process 121](#_Toc35407693)

[10 Parsing process 125](#_Toc35407694)

[10.1 General 125](#_Toc35407695)

[10.2 Parsing process for 0-th order Exp-Golomb codes 126](#_Toc35407696)

[10.2.1 General 126](#_Toc35407697)

[10.2.2 Mapping process for signed Exp-Golomb codes 127](#_Toc35407698)

[Annex A Profiles, tiers, and levels 129](#_Toc35407699)

[A.1 Overview of profiles, tiers, and levels 129](#_Toc35407700)

[A.2 V-PCC profile, tier and level structure 129](#_Toc35407701)

[A.3 V-PCC CodecGroup profile components 132](#_Toc35407702)

[A.4 V-PCC toolset profile components 133](#_Toc35407703)

[A.4.1 Basic and Extended toolset profile component 133](#_Toc35407704)

[A.5 V-PCC reconstruction profile components 134](#_Toc35407706)

[A.5.1 V-PCC reconstruction decoder conformance 135](#_Toc35407707)

[A.5.1.1 Hypothetical reference chroma format conversion 135](#_Toc35407708)

[A.5.1.2 Hypothetical reference resolution upscaling 135](#_Toc35407709)

[A.5.1.3 Hypothetical reference frame rate conversion 136](#_Toc35407710)

[A.5.1.4 Hypothetical reference attribute transfer process 136](#_Toc35407711)

[A.6 Tiers and Levels 136](#_Toc35407712)

[Annex B Post-decoding conversion to nominal video formats 140](#_Toc35407713)

[B.1 Nominal video formats 140](#_Toc35407714)

[B.2 Video format conversion process 140](#_Toc35407715)

[B.2.1 General process 140](#_Toc35407716)

[B.2.2 Conversion of a two dimensional array to a nominal bitdepth 140](#_Toc35407717)

[B.2.3 Video format conversion process for resolution 141](#_Toc35407718)

[B.2.4 Video format conversion process for chroma upsampling 141](#_Toc35407719)

[B.2.5 Video format conversion process for frame rate 141](#_Toc35407720)

[B.2.6 Video format conversion process for attribute transfer 141](#_Toc35407721)

[B.3 Geometry compositing process 141](#_Toc35407722)

[B.4 Attribute compositing process 142](#_Toc35407723)

[B.5 Attribute dimension packing process for an attribute with attrIdx 144](#_Toc35407724)

[Annex C V-PCC Sample stream format 146](#_Toc35407725)

[C.1 General 146](#_Toc35407726)

[C.2 Sample stream V-PCC unit syntax and semantics 146](#_Toc35407727)

[C.2.1 Sample stream V-PCC header syntax 146](#_Toc35407728)

[C.2.2 Sample stream V-PCC unit syntax 146](#_Toc35407729)

[C.2.3 Sample stream V-PCC header semantics 146](#_Toc35407730)

[C.2.4 Sample stream V-PCC unit semantics 147](#_Toc35407731)

[Annex D NAL Sample stream format 148](#_Toc35407732)

[D.1 General 148](#_Toc35407733)

[D.2 Sample stream NAL unit syntax and semantics 148](#_Toc35407734)

[D.2.1 Sample stream NAL header syntax 148](#_Toc35407735)

[D.2.2 Sample stream NAL unit syntax 148](#_Toc35407736)

[D.2.3 Sample stream NAL header semantics 148](#_Toc35407737)

[D.2.4 Sample stream NAL unit semantics 149](#_Toc35407738)

[Annex E Atlas Hypothetical Reference Decoder 150](#_Toc35407739)

[E.1 General 150](#_Toc35407740)

[E.2 Operation of coded atlas frame buffer 154](#_Toc35407741)

[E.2.1 General 154](#_Toc35407742)

[E.2.2 Timing of decoding unit arrival 155](#_Toc35407743)

[E.2.3 Timing of decoding unit removal and decoding of decoding unit 157](#_Toc35407744)

[E.3 Operation of the decoded atlas frame buffer 158](#_Toc35407745)

[E.3.1 General 158](#_Toc35407746)

[E.3.2 Removal of atlas frames from the DAB before decoding of the current atlas frame 159](#_Toc35407747)

[E.3.3 Atlas frame output 159](#_Toc35407748)

[E.3.4 Current decoded atlas frame marking and storage 160](#_Toc35407749)

[E.3.5 Removal of atlas frames from the DAB after decoding of the current atlas frame 160](#_Toc35407750)

[E.4 Bitstream conformance 160](#_Toc35407751)

[E.5 Decoder conformance 162](#_Toc35407752)

[E.5.1 General 162](#_Toc35407753)

[E.5.2 Operation of the output order DAB 163](#_Toc35407754)

[E.5.2.1 General 163](#_Toc35407755)

[E.5.2.2 Output and removal of atlas frames from the DAB 163](#_Toc35407756)

[E.5.2.3 Additional bumping 164](#_Toc35407757)

[E.5.2.4 "Bumping" process 165](#_Toc35407758)

[Annex F Supplemental enhancement information 166](#_Toc35407759)

[F.1 General 166](#_Toc35407760)

[F.2 SEI payload syntax 166](#_Toc35407761)

[F.2.1 General SEI message syntax 166](#_Toc35407762)

[F.2.2 Filler payload SEI message syntax 168](#_Toc35407763)

[F.2.3 User data registered by Recommendation ITU-T T.35 SEI message syntax 168](#_Toc35407764)

[F.2.4 User data unregistered SEI message syntax 168](#_Toc35407765)

[F.2.5 Recovery point SEI message syntax 169](#_Toc35407766)

[F.2.6 No display SEI message syntax 169](#_Toc35407767)

[F.2.7 Reserved SEI message syntax 169](#_Toc35407768)

[F.2.8 SEI manifest SEI message syntax 169](#_Toc35407769)

[F.2.9 SEI prefix indication SEI message syntax 169](#_Toc35407770)

[F.2.10 Attribute transformation parameters SEI message syntax 170](#_Toc35407771)

[F.2.11 Active sub-bitstreams SEI message syntax 171](#_Toc35407772)

[F.2.12 Component codec mapping SEI message syntax 171](#_Toc35407773)

[F.2.13 Volumetric Annotation SEI message family syntax 171](#_Toc35407774)

[F.2.13.1 Scene Object Information SEI message syntax 171](#_Toc35407775)

[F.2.13.2 Object Label Information SEI message syntax 174](#_Toc35407776)

[F.2.13.3 Patch Information SEI message syntax 174](#_Toc35407777)

[F.2.13.4 Volumetric rectangle information SEI message syntax 175](#_Toc35408055)

[F.2.14 Buffering period SEI message syntax 177](#_Toc35408056)

[F.2.15 Atlas frame timing SEI message syntax 177](#_Toc35408057)

[F.2.16 Presentation information SEI message syntax 177](#_Toc35408058)

[F.2.17 Geometry smoothing SEI message syntax 178](#_Toc35408059)

[F.2.18 Attribute smoothing SEI message syntax 179](#_Toc35408060)

[F.3 SEI payload semantics 180](#_Toc35408061)

[F.3.1 General SEI payload semantics 180](#_Toc35408062)

[F.3.2 Filler payload SEI message semantics 181](#_Toc35408063)

[F.3.3 User data registered by Recommendation ITU-T T.35 SEI message semantics 181](#_Toc35408064)

[F.3.4 User data unregistered SEI message semantics 182](#_Toc35408065)

[F.3.5 Recovery point SEI message semantics 182](#_Toc35408066)

[F.3.6 No display SEI message semantics 183](#_Toc35408067)

[F.3.7 Reserved SEI message semantics 183](#_Toc35408068)

[F.3.8 SEI manifest SEI message semantics 183](#_Toc35408069)

[F.3.9 SEI prefix indication SEI message semantics 185](#_Toc35408070)

[F.3.10 Attribute transformation parameters SEI message semantics 186](#_Toc35408098)

[F.3.11 Active sub-bitstreams SEI message semantics 186](#_Toc35408099)

[F.3.12 Component codec mapping SEI message semantics 187](#_Toc35408100)

[F.3.13 Volumetric annotation SEI messages family syntax 188](#_Toc35408101)

[F.3.13.1 Scene Object information SEI message semantics 188](#_Toc35408102)

[F.3.13.2 Object label information SEI message semantics 191](#_Toc35408103)

[F.3.13.3 Patch information SEI message semantics 192](#_Toc35408104)

[F.3.13.4 Volumetric rectangle information SEI message semantics 193](#_Toc35408105)

[F.3.14 Buffering period SEI message semantics 194](#_Toc35408106)

[F.3.15 Atlas frame timing SEI message semantics 196](#_Toc35408107)

[F.3.16 Presentation information SEI message semantics 198](#_Toc35408108)

[F.3.17 Geometry smoothing SEI message semantics 199](#_Toc35408109)

[F.3.18 Attribute smoothing SEI message semantics 200](#_Toc35408110)

[Annex G Volumetric usability information 203](#_Toc35408111)

[G.1 General 203](#_Toc35408112)

[G.2 VUI syntax 203](#_Toc35408113)

[G.2.1 VUI parameters syntax 203](#_Toc35408114)

[G.2.2 HRD parameters syntax 203](#_Toc35408115)

[G.2.3 Sub-layer HRD parameters syntax 204](#_Toc35408116)

[G.3 VUI semantics 204](#_Toc35408117)

[G.3.1 VUI parameters semantics 204](#_Toc35408118)

[G.3.2 HRD parameters semantics 205](#_Toc35408119)

[G.3.3 HRD sub-layer parameters semantics 207](#_Toc35408120)

[Bibliography 209](#_Toc35408121)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO documents 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](https://www.iso.org/directives-and-policies.html)).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO 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](https://www.iso.org/iso-standards-and-patents.html)).

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](https://www.iso.org/foreword-supplementary-information.html).

This document was prepared by Subcommittee 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](https://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. Point clouds have arisen as one of the main representations for such applications. A point cloud frame consists of a set of individual 3D points. Each point, in addition to having a 3D position, i.e., spatial attribute, may also be associated with a number of other attributes such as colour, reflectance, surface normal, etc. A point cloud consists of a sequence of point cloud frames. The number of points, their positions, and their attributes may vary from one frame to another. Such representations require a large amount of data, which can be costly in terms of storage and transmission. Therefore, the ISO/IEC Moving Picture Experts Group (MPEG) developed a new Recommendation | International Standard, which aims at efficiently compressing point cloud representations.

**Information technology — Coded Representation of Immersive Media — Part 5: Video-based Point Cloud Compression**

# Scope

This document specifies video-based point cloud compression.

# Normative reference

There are no normative references in this document.

[Ed. (KM): Revisit. Should AVC and HEVC be added in this section?]

# Terms and definitions

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

[Ed. (All): Ensure that all the definitions are useful. Some definitions for example are only used once or twice inside the document. There could also be missing definitions that need to be defined in the document.]

associated non-ACL NAL unit

a *non-ACL NAL unit* (when present) for an *ACL NAL unit* where the *ACL NAL unit* is the *associated ACL NAL unit* of the *non-ACL NAL unit*

associated ACL NAL unit

the preceding *ACL NAL unit* in *decoding order* for a *non-ACL NAL unit* with nal\_unit\_type equal to NAL\_EOS, NAL\_EOB, NAL\_FD, NAL\_SUFFIX\_NSEI, or NAL\_SUFFIX\_ESEI, or in the ranges of NAL\_RSV\_NACL\_45..NAL\_RSV\_NACL\_47 or NAL\_UNSPEC\_56..NAL\_UNSPEC\_63; or otherwise the next *ACL NAL unit* in *decoding order*

atlas

a collection of 2D bounding boxes, i.e. patches, projected into a rectangular frame that correspond to a 3-dimensional bounding box in 3D space, which may represent a subset of a point cloud

atlas frame

See atlas

atlas frame parameter set (AFPS)

a syntax structure containing syntax elements that apply to zero or more entire coded atlas frames as determined by the content of a syntax element found in the tile group header

atlas sequence

a collection of atlas frames



atlas sequence parameter set (ASPS)

a syntax structure containing syntax elements that apply to zero or more entire coded atlas sequences (CASs) as determined by the content of a syntax element found in the ASPS referred to by a syntax element found in each tile group header

[Ed. LK: can we consider changing the naming to sequence parameter set, frame parameter set without the prefix atlas. The sequence parameter set on vpcc unit level could be re-named to video parameter set (this would be aligned with HEVC)]

[Ed. (ALL): preliminary discussion to keep it as is since it appears to clarify that we are referring to atlases. Should be revisited during the upcoming MPEG meeting]

atlas sub-bitstream

an extracted sub-bitstream from the V-PCC bitstream containing a part of an atlas *NAL bitstream*

attribute

scalar or vector property optionally associated with each point in a point cloud such as colour, reflectance, surface normal, time stamps, material ID, etc

attribute frame

a collection of attribute maps for a specific attribute that correspond to the same time instance

attribute map

a 2D rectangular array created through the aggregation of *patches* with *attribute* values for a specific attribute

bitstream

A sequence of bits of coded *V-PCC components*

byte

a 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

byte-aligned

a position in a *bitstream* is byte-aligned when the position is an integer multiple of 8 bits from the position of the first bit in the *bitstream*, and a bit or *byte* or *syntax element* is said to be byte-aligned when the position at which it appears in a *bitstream* is byte-aligned

cartesian coordinates

three scalars (x, y, z) with finite precision and dynamic range that indicate the location of a point relative to a fixed reference point (the origin)

coded atlas

a coded representation of an *atlas*

coded atlas access unit

a set of atlas *nal units* that are associated with each other according to a specified classification rule, are consecutive in *decoding order*, and contain all atlas *nal\_units* pertaining to one particular output time

coded atlas bitstream

sequence of bits that forms the representation of *atlas frames* and associated data forming one or more *CASs*

coded atlas frame

see *coded* *atlas*

coded atlas sequence (CAS)

a sequence of *coded atlas access units,* in decoding order, of an IRAP *coded atlas* AU, followed by zero or more *coded atlas* AUs that are not IRAP *coded atlas* AUs, including all subsequent *coded atlas* AUs up to but not including any subsequent *coded atlas* AU that is an IRAP *coded atlas* AU

coded point cloud frame

a collection of coded representations of an *atlas*, *occupancy map*, one or more *geometry maps*, and, for each available attribute, one or more *attribute maps*,pertaining to one particular time instance

coded point cloud sequence (CPCS)

sequence of V-PCC AUs that consists, in decoding order, of an IRAP V-PCC AU, followed by zero or more V-PCC AUs that are not IRAP V-PCC AUs, including all subsequent V-PCC AUs up to but not including any subsequent V-PCC AU that is an IRAP V-PCC AU

[Ed. (YK): Can V-PCC units within a CPCS refer to different V-PCC VPSs? If yes, then what does the description "sequence level parameters" in Table 7-1 mean? If not, a few syntax elements, e.g., vps\_map\_count\_minus1, would apply to all maps of geometry and, at the same time, all maps of all types of attribute! Some change is needed, e.g., making such information specific to each sub-bitstream.]

[Ed. (All): We agree that this definition is problematic since it does not handle the case where the different components are coded independently.]

coded representation

a data element as represented in its coded form

coded V-PCC access unit

a coded representation of a V-PCC access unit

coded V-PCC component

a coded representation of a V-PCC component

collection of V-PCC sub-bitstream components

V-PCC sub-bitstream components that when decoded, they enable the reconstruction of a point cloud

essential supplemental enhancement information (ESEI)

a NAL unit corresponding to an essential SEI and has nal\_unit\_type equal to NAL\_PREFIX\_ESEI or NAL\_SUFFIX\_ESEI

flag

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

geometry

set of cartesian coordinates associated with a point cloud frame

geometry frame

a collection of geometry maps corresponding to the same time instance

geometry map

a 2D array created through the aggregation of the geometry information associated with each *patch*

instantaneous decoding refresh (IDR) coded atlas access unit

an access unit in which the coded atlas with nal\_layer\_id equal to 0 is an IDR or a GIDR coded atlas.

instantaneous decoding refresh (IDR) coded atlas

an IRAP coded atlas for which each ACL NAL unit has nal\_unit\_type equal to NAL\_IDR\_W\_RADL, or NAL\_IDR\_N\_LP, NAL\_GIDR\_W\_RADL, or NAL\_GIDR\_N\_LP.

NOTE – An IDR coded atlas does not refer to any atlases other than itself for inter prediction in its decoding process, and may be the first atlas in the bitstream in decoding order, or may appear later in the bitstream. Each IDR coded atlas is the first atlas of a CAS in decoding order. When an IDR coded atlas for which each ACL NAL unit has nal\_unit\_type equal to NAL\_IDR\_W\_RADL or NAL\_GIDR\_W\_RADL, it may have associated RADL coded atlases. When an IDR coded atlas for which each ACL NAL unit has nal\_unit\_type equal to NAL\_IDR\_N\_LP or NAL\_GIDR\_N\_LP, it does not have any associated leading coded atlases. An IDR coded atlas does not have associated RASL coded atlases.

instantaneous decoding refresh (IDR) V-PCC access unit

an access unit in which the coded atlas with nal\_layer\_id equal to 0 is an GIDR coded atlas

instantaneous decoding refresh (IDR) V-PCC

an IRAP coded atlas for which each ACL NAL unit has nal\_unit\_type equal to NAL\_GIDR\_W\_RADL or NAL\_GIDR\_N\_LP

NOTE – A GIDR coded atlas does not refer to any atlases other than itself for inter prediction in its decoding process, and may be the first atlas in the bitstream in decoding order, or may appear later in the bitstream. Each GIDR coded atlas is the first atlas of a CAS in decoding order. When an IDR coded atlas for which each ACL NAL unit has nal\_unit\_type equal to NAL\_IDR\_W\_RADL, it may have associated RADL coded atlases. When an IDR coded atlas for which each ACL NAL unit has nal\_unit\_type equal to NAL\_IDR\_N\_LP, it does not have any associated leading coded atlases. An IDR coded atlas does not have associated RASL coded atlases.

inter atlas tile group

an atlas tile group that is decoded using both intra or inter prediction methods

intra atlas tile group

an atlas tile group that is decoded using only intra prediction methods



intra random access point (IRAP) coded atlas

A coded atlas for which each ACL NAL unit has nal\_unit\_type in the range of NAL\_BLA\_W\_LP to NAL\_IRAP\_ACL\_23, inclusive

NOTE – An IRAP coded atlas does not refer to any coded atlases other than itself for inter prediction in its decoding process, and may be a BLA coded atlas, GBLA coded atlas, a CRA coded atlas, a GCRA coded atlas, an IDR coded atlas, or a GIDR coded atlas. The first coded atlas in the bitstream in decoding order must be an IRAP coded atlas. Provided the necessary parameter sets are available when they need to be activated, the IRAP coded atlas and all subsequent non-RASL coded atlas in decoding order can be correctly decoded without performing the decoding process of any coded atlases that precede the IRAP coded atlas in decoding order.

intra random access point (IRAP) coded atlas access unit

An access unit in which the coded atlas with nal\_layer\_id equal to 0 is an IRAP coded atlas

intra random access point (IRAP) V-PCC unit

A V-PCC unit for which each ACL NAL unit has nal\_unit\_type in the range of NAL\_GBLA\_W\_LP to NAL\_GBLA\_N\_LP, or is in the range of NAL\_GIDR\_W\_RADL to NAL\_GIDR\_N\_LP, or is equal to NAL\_GCRA, inclusive

intra random access point (IRAP) V-PCC access unit

An access unit that corresponds to an IRAP V-PCC unit

network abstraction layer (NAL) unit

syntax structure containing an indication of the type of data to follow and bytes containing that data in the form of an RBSP

network abstraction layer (NAL) unit stream

sequence of *NAL units*

non-essential supplemental enhancement information (NSEI)

a NAL unit corresponding to a non-essential SEI and has nal\_unit\_type equal to NAL\_PREFIX\_NSEI or NAL\_SUFFIX\_NSEI

occupancy map

a 2D array corresponding to an atlas whose values indicate for each sample position in the atlas whether that position corresponds to a valid 3D point in the point cloud representation

patch

rectangular region within an *atlas* that corresponds to a rectangular region within a *planar projection*

[Ed. [LK] This definition has to be refined. Originated from m48101]

patch data

data that is needed to perform the transformation of patches included in an atlas from 2D to 3D space

planar projection

2D sample arrays of *attributes* and a corresponding *geometry* representing the projection of a *point cloud frame* onto a planar surface

[Ed. [All] : Revisit]

point cloud

set of 3D points specified by their cartesian coordinates (x,y,z) and zero or more corresponding sets of attributes at a particular time instance

point cloud frame

see *point cloud*

point cloud sequence

sequence of *point cloud frames*

prefix ESEI atlas NAL unit

an essential SEI atlas NAL unit that has nal\_unit\_type equal to NAL\_PREFIX\_ESEI

prefix NSEI atlas NAL unit

a non-essential SEI atlas NAL unit that has nal\_unit\_type equal to NAL\_PREFIX\_NSEI

prefix SEI message

an SEI message that is contained in a prefix NSEI or prefix ESEI atlas NAL unit

prefix SEI atlas NAL unit

an SEI atlas NAL unit that has nal\_unit\_type equal to NAL\_PREFIX\_NSEI or NAL\_PREFIX\_ESEI

random access

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



raw byte sequence payload (RBSP)

a *syntax structure* containing an integer number of *bytes* that is encapsulated in a *NAL unit* and that is either empty or has the form of a *string of data bits* containing *syntax elements* followed by an *RBSP stop bit* and zero or more subsequent bits equal to 0

raw byte sequence payload (RBSP) stop bit

A bit equal to 1 present within a *raw byte sequence payload (RBSP)* after a *string of data bits*, for which the location of the end within an *RBSP* can be identified by searching from the end of the *RBSP* for the *RBSP stop bit*, which is the last non-zero bit in the *RBSP*

v-pcc parameter set (VPS)

a syntax structure containing syntax elements that apply to zero or more entire CPCSs as determined by the content of a syntax element found in the VPS referred to by a syntax element found in the V-PCC unit header

suffix ESEI NAL unit

an essential SEI atlas nal unit that has nal\_unit\_type equal to NAL\_SUFFIX\_ESEI

suffix NSEI NAL unit

a non-essential SEI atlas nal unit that has nal\_unit\_type equal to NAL\_SUFFIX\_NSEI

suffix SEI message

an SEI message that is contained in a suffix NSEI or ESEI NAL unit

suffix SEI NAL unit

an SEI atlas nal unit that has nal\_unit\_type equal to NAL\_SUFFIX\_NSEI or NAL\_SUFFIX\_ESEI

supplemental enhancement information (SEI)

a NAL unit that has nal\_unit\_type equal to NAL\_PREFIX\_NSEI, NAL\_PREFIX\_ESEI, NAL\_SUFFIX\_NSEI, or NAL\_SUFFIX\_ESEI

[Ed. AMT: there is a question here if this may be confused to NAL units specified in the video codec. Maybe we need to clarify that NAL units here correspond to the units specified in this document only.]

syntax element

an element of data represented in the bitstream

syntax structure

zero or more syntax elements present together in the bitstream in a specified order

tile

a rectangular region of an *atlas*

[Ed. [All] to be revisited]

tile group

a group of *tiles* ...

[Ed. [All] It has to be refined]

video bitstream

a bitstream conforming to a video specification that may represent a V-PCC component as specified by this document

[Ed. (ALL): can a single NAL unit, that may refer to other NAL units in the bitstream, be considered as a video bitstream on its own? Or does the video bitstream need to be "self contained"? This impacts the definition of the video sub-bitstream, and also relates to the definition of the atlas sub-bitstream]

video data unit

syntax structure containing video data information, such as a NAL unit in the context of HEVC or AVC

video sub-bitstream

an extracted sub-bitstream from the V-PCC bitstream containing a part of a *video bitstream*

[Ed. (ALL): There may be some ambiguity with what a "part of a video bitstream" may mean. We should not allow for example splitting a NAL unit into smaller parts]

V-PCC access unit

a set of *V-PCC units* that are associated with each other according to a specified classification rule, and are consecutive in *decoding order*, and contain all atlas nal\_units as well as other V-PCC components pertaining to one particular output time

V-PCC bitstream

sequence of bits that forms the representation of *coded point cloud frames* and associated data forming one or more *CPCSs*

V-PCC component

an *atlas*, *occupancy map*, *geometry,* or *attribute* of a particular type that is associated with a V-PCC point cloud representation

V-PCC component frame

a single V-PCC component pertaining to one particular output time

V-PCC sub-bitstream

a sub-bitstream of a V-PCC bitstream corresponding to a V-PCC component

V-PCC unit

a *syntax structure* containing an *V-PCC unit header* and a *V-PCC unit payload*

V-PCC unit header

a *syntax structure* containing an indication of the type of data to follow

V-PCC unit payload

a *syntax structure* containing the *bytes* containing V-PCC sub-bitstream data

[Ed. (All): some definitions are inconsistent; additional definitions may be necessary]

# Abbreviations

For the purposes of this International Standard, the following abbreviations apply:

2D Two-Dimensional

3D Three-Dimensional

ACL Atlas Coding Layer

AAPS Atlas Adaptation Parameter Set

AFPS Atlas Frame Parameter Set

ASPS Atlas Sequence Parameter Set

AU Access Unit

CAS Coded Atlas Sequence

CPCS Coded Point Cloud Sequence

DAFB Decoded Atlas Frame Buffer

EOM Enhanced Occupancy Map

I Intra

IDR Instantaneous Decoding Refresh

IRAP Intra Random Access Point

LSB Least Significant Bit

MSB Most Significant Bit

NAL Network Abstraction Layer

P Predictive

PCC Point Cloud Compression

PLR Point Local Reconstruction

RBSP Raw Byte Sequence Payload

SEI Supplemental Enhancement Information

SODB String of Data Bits

VPS V-PCC Parameter Set

V-PCC Video-based PCC

# Conventions

## General

NOTE – The mathematical operators used in this document are similar to those used in the C programming language. However, the results of integer division and arithmetic shift operations are defined more precisely, and additional operations are defined, such as exponentiation and real-valued division. Numbering and counting conventions generally begin from 0.

## Arithmetic operators

The following arithmetic operators are defined as follows:

|  |  |
| --- | --- |
| + | Addition |
| − | Subtraction (as a two-argument operator) or negation (as a unary prefix operator) |
| \* | Multiplication, including matrix multiplication |
| xy | Exponentiation. Specifies x to the power of y. In other contexts, such notation is used for superscripting not intended for interpretation as exponentiation. |
| / | Integer division with truncation of the result toward zero. For example, 7 / 4 and −7 / −4 are truncated to 1 and −7 / 4 and 7 / −4 are truncated to −1. |
| ÷ | Used to denote division in mathematical equations where no truncation or rounding is intended. |
|  | Used to denote division in mathematical equations where no truncation or rounding is intended. |
|  | The summation of f( i ) with i taking all integer values from x up to and including y. |
| x % y | Modulus. Remainder of x divided by y, defined only for integers x and y with x >= 0 and y > 0. |

## Logical operators

The following logical operators are defined as follows:

|  |  |
| --- | --- |
| x && y | Boolean logical "and" of x and y. |
| x  | |  y | Boolean logical "or" of x and y. |
| ! | Boolean logical "not". |
| x ? y : z | If x is TRUE or not equal to 0, evaluates to the value of y; otherwise, evaluates to the value of z. |

## Relational operators

The following relational operators are defined as follows:

|  |  |
| --- | --- |
| > | Greater than. |
| >= | Greater than or equal to. |
| < | Less than. |
| <= | Less than or equal to. |
| = = | Equal to. |
| != | Not equal to. |

## Bit-wise operators

The following bit-wise operators are defined as follows:

|  |  |
| --- | --- |
| ~ | Bit-wise "not". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| & | Bit-wise "and". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| | | Bit-wise "or". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| ^ | Bit-wise "exclusive or". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0. |
| x >> y | Arithmetic right shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values of y. Bits shifted into the MSBs as a result of the right shift have a value equal to the MSB of x prior to the shift operation. |
| x << y | Arithmetic left shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values of y. Bits shifted into the LSBs as a result of the left shift have a value equal to 0. |

## Assignment operators

The following arithmetic operators are defined as follows:

|  |  |
| --- | --- |
| = | Assignment operator. |
| + + | Increment, i.e. x+ + is equivalent to x = x + 1; when used in an array index, evaluates to the value of the variable prior to the increment operation. |
| − − | Decrement, i.e. x− − is equivalent to x = x − 1; when used in an array index, evaluates to the value of the variable prior to the decrement operation. |
| += | Increment by amount specified, i.e. x += 3 is equivalent to x = x + 3, and x += (−3) is equivalent to x = x + (−3). |
| −= | Decrement by amount specified, i.e. x −= 3 is equivalent to x = x − 3, and x −= (−3) is equivalent to x = x − (−3). |

## Range notation

The following notation is used to specify a range of values:

|  |  |
| --- | --- |
| x = y..z | x takes on integer values starting from y to z, inclusive, with x, y, and z being integer numbers and z being greater than y. |

## Mathematical functions

The following mathematical functions are defined:

Abs( x ) = (5‑1)

Ceil( x ) the smallest integer greater than or equal to x. (5‑2)

Clip1Y( x ) = Clip3( 0, ( 1 << BitDepthY ) – 1, x ) (5‑3)

Clip1C( x ) = Clip3( 0, ( 1 << BitDepthC ) – 1, x ) (5‑4)

Clip3( x, y, z ) = (5‑5)

Floor( x ) the largest integer less than or equal to x. (5‑6)

Log2( x ) the base-2 logarithm of x. (5‑7)

Log10( x ) the base-10 logarithm of x. (5‑8)

Min( x, y ) = (5‑9)

Max( x, y ) = (5‑10)

Round( x ) = Sign( x ) \* Floor( Abs( x ) + 0.5 ) (5‑11)

Sign( x ) = (5‑12)

Sqrt( x ) the square root of x (5‑13)

[Ed. (All): add missing definitions from HEVC or CICP specs; remove unused definitions.]

## Order of operation precedence

When order of precedence in an expression is not indicated explicitly by use of parentheses, the following rules apply:

– Operations of a higher precedence are evaluated before any operation of a lower precedence.

– Operations of the same precedence are evaluated sequentially from left to right.

Table 5‑1 specifies the precedence of operations from highest to lowest; a higher position in the table indicates a higher precedence.

NOTE – For those operators that are also used in the C programming language, the order of precedence used in this document is the same as used in the C programming language.

Table 5‑1 – Operation precedence from highest (at top of table) to lowest (at bottom of table)

|  |  |
| --- | --- |
| **operations (with operands x, y, and z)** | |
| "x++", "x− −" |  |
| "!x", "−x" (as a unary prefix operator) |  |
| xy |  |
| "x \* y", "x / y", "x ÷ y", "", "x % y" |  |
| "x + y", "x − y" (as a two-argument operator), |  |
| "x  <<  y", "x  >>  y" |  |
| "x < y", "x  <=  y", "x > y", "x  >=  y" |  |
| "x  = =  y", "x  !=  y" |  |
| "x & y" |  |
| "x | y" |  |
| "x  &&  y" |  |
| "x  | |  y" |  |
| "x ? y : z" |  |
| "x..y" |  |
| "x = y", "x  +=  y", "x  −=  y" |  |

[Ed. (All): fix table according to HEVC; some elements seem to be missing.]

## Variables, syntax elements, and tables

Syntax elements in the bitstream are represented in **bold** type. Each syntax element is described by its name (all lower case letters with underscore characters), and one descriptor for its method of coded representation. The decoding process behaves according to the value of the syntax element and to the values of previously decoded syntax elements. When a value of a syntax element is used in the syntax tables or the text, it appears in regular (i.e. not bold) type.

In some cases the syntax tables may use the values of other variables derived from syntax elements values. Such variables appear in the syntax tables, or text, named by a mixture of lower case and upper case letter and without any underscore characters. Variables starting with an upper case letter are derived for the decoding of the current syntax structure and all depending syntax structures. Variables starting with an upper case letter may be used in the decoding process for later syntax structures without mentioning the originating syntax structure of the variable. Variables starting with a lower case letter are only used within the subclause in which they are derived.

In some cases, "mnemonic" names for syntax element values or variable values are used interchangeably with their numerical values. Sometimes "mnemonic" names are used without any associated numerical values. The association of values and names is specified in the text. The names are constructed from one or more groups of letters separated by an underscore character. Each group starts with an upper case letter and may contain more upper case letters.

NOTE 1 – The syntax is described in a manner that closely follows the C-language syntactic constructs.

Functions that specify properties of the current position in the bitstream are referred to as syntax functions. These functions are specified in clause 7.2 and assume the existence of a bitstream pointer with an indication of the position of the next bit to be read by the decoding process from the bitstream. Syntax functions are described by their names, which are constructed as syntax element names and end with left and right round parentheses including zero or more variable names (for definition) or values (for usage), separated by commas (if more than one variable).

Functions that are not syntax functions (including mathematical functions specified in clause 5.8) are described by their names, which start with an upper case letter, contain a mixture of lower and upper case letters without any underscore character, and end with left and right parentheses including zero or more variable names (for definition) or values (for usage) separated by commas (if more than one variable).

A one-dimensional array is referred to as a list. A two-dimensional array is referred to as a matrix. Arrays can either be syntax elements or variables. Subscripts or square parentheses are used for the indexing of arrays. In reference to a visual depiction of a matrix, the first subscript is used as a row (vertical) index and the second subscript is used as a column (horizontal) index. The same indexing order is used when using square parentheses. Thus, an element of a matrix s at horizontal position x and vertical position y may be denoted either as s[ y ][ x ] or as syx. A single row of a matrix may be referred to as a list and denoted by omission of the column index. Thus, the row of a matrix s at horizontal position x may be referred to as the list s[ y ].

NOTE 2 – In some video specifications a reverse indexing order may be used when using square parenthesis for depicting two dimensional arrays, i.e. the first element in the square parentheses is the column (horizontal) index and the second element in the square parentheses is the row (vertical) index.

A specification of values of the entries in rows and columns of an array may be denoted by { {...} {...} }, where each inner pair of brackets specifies the values of the elements within a row in increasing column order and the rows are ordered in increasing row order. Thus, setting a matrix s equal to { { 1 6 } { 4 9 } specifies that s[ 0 ][ 0 ] is set equal to 1, s[ 0 ][ 1 ] is set equal to 6, s[ 1 ][ 0 ] is set equal to 4, and s[ 1 ][ 1 ] is set equal to 9.

Binary notation is indicated by enclosing the string of bit values by single quote marks. For example, '01000001' represents an eight-bit string having only its second and its last bits (counted from the most to the least significant bit) equal to 1.

Hexadecimal notation, indicated by prefixing the hexadecimal number by "0x", may be used instead of binary notation when the number of bits is an integer multiple of 4. For example, 0x41 represents an eight-bit string having only its second and its last bits (counted from the most to the least significant bit) equal to 1.

Numerical values not enclosed in single quotes and not prefixed by "0x" are decimal values.

A value equal to 0 represents a FALSE condition in a test statement. The value TRUE is represented by any value different from zero.

## Text description of logical operations

In the text, a statement of logical operations as would be described mathematically in the following form:

if( condition 0 )  
 statement 0  
else if( condition 1 )  
 statement 1  
...  
else /\* informative remark on remaining condition \*/  
 statement n

may be described in the following manner:

... as follows / ... the following applies:

– If condition 0, statement 0

– Otherwise, if condition 1, statement 1

– ...

– Otherwise (informative remark on remaining condition), statement n

Each "If ... Otherwise, if ... Otherwise, ..." statement in the text is introduced with "... as follows" or "... the following applies" immediately followed by "If ... ". The last condition of the "If ... Otherwise, if ... Otherwise, ..." is always an "Otherwise, ...". Interleaved "If ... Otherwise, if ... Otherwise, ..." statements can be identified by matching "... as follows" or "... the following applies" with the ending "Otherwise, ...".

In the text, a statement of logical operations as would be described mathematically in the following form:

if( condition 0a && condition 0b )  
 statement 0  
else if( condition 1a | | condition 1b )  
 statement 1  
...  
else  
 statement n

may be described in the following manner:

... as follows / ... the following applies:

– If all of the following conditions are true, statement 0:

– condition 0a

– condition 0b

– Otherwise, if one or more of the following conditions are true, statement 1:

– condition 1a

– condition 1b

– ...

– Otherwise, statement n

In the text, a statement of logical operations as would be described mathematically in the following form:

if( condition 0 )  
 statement 0  
if( condition 1 )  
 statement 1

may be described in the following manner:

When condition 0, statement 0

When condition 1, statement 1

## Processes

Processes are used to describe the decoding of syntax elements. A process has a separate specification and invoking. All syntax elements and upper case variables that pertain to the current syntax structure and depending syntax structures are available in the process specification and invoking. A process specification may also have a lower case variable explicitly specified as input. Each process specification has explicitly specified an output. The output is a variable that can either be an upper case variable or a lower case variable.

When invoking a process, the assignment of variables is specified as follows:

– If the variables at the invoking and the process specification do not have the same name, the variables are explicitly assigned to lower case input or output variables of the process specification.

– Otherwise (the variables at the invoking and the process specification have the same name), assignment is implied.

In the specification of a process, a specific coding block may be referred to by the variable name having a value equal to the address of the specific coding block.

# Bitstream format, partitioning, and scanning processes

## V-PCC bitstream formats

This clause specifies the relationship between the V-PCC unit stream format and the V-PCC sample stream, either of which are referred to as the V-PCC bitstream. This clause is not an essential component of this document and all V-PCC components including any associated V-PCC VPSs could be encapsulated using a different format depending on application.

The bitstream can be in one of two formats: the V-PCC unit stream format or the sample stream format. The V-PCC unit stream format is conceptually the more "basic" type. It consists of a sequence of syntax structures called V-PCC units. This sequence is ordered in decoding order. There are constraints imposed on the decoding order (and contents) of the V-PCC units in the V-PCC unit stream.

NOTE – The V-PCC unit stream format is commonly not intended to be used in any applications on its own since it requires additional information, i.e. sub-bitstream size information, for decoding its associated sub-bitstreams. One method of achieving this is through the use of the sample stream format.

The sample stream format can be constructed from the V-PCC unit stream format by ordering the V-PCC units in decoding order and prefixing each V-PCC unit with a heading that specifies the exact size, in bytes, of the V-PCC unit. A sample stream header is included at the beginning of the sample stream bitstream that specifies the precision, in bytes, of the signaled V-PCC unit size. The V-PCC unit stream format can be extracted from the sample stream format by traversing through the sample stream format, reading the size information and appropriately extracting each V-PCC unit. Methods of framing V-PCC units in a manner other than the use of the sample stream format are outside the scope of this document. The sample stream format is specified in Annex B.

## NAL bitstream formats

This clause specifies the relationship between the network abstraction layer (NAL) unit stream and the NAL sample stream, either of which are referred to as the NAL bitstream.

The bitstream can be in one of two formats: the NAL unit stream format or the sample stream format. The NAL unit stream format is conceptually the more "basic" type. It consists of a sequence of syntax structures called NAL units. This sequence is ordered in decoding order, as described in clause 7.4.5.3. There are constraints imposed on the decoding order (and contents) of the NAL units in the NAL unit stream.

NOTE – The NAL unit stream format is commonly not intended to be used in any applications on its own since it requires additional information, i.e. sub-bitstream size information, for decoding its associated sub-bitstreams. One method of achieving this is through the use of the NAL sample stream format.

The NAL sample stream format can be constructed from the NAL unit stream format by ordering the NAL units in decoding order and prefixing each NAL unit with a heading that specifies the exact size, in bytes, of the NAL unit. A sample stream header is included at the beginning of the sample stream bitstream that specifies the precision, in bytes, of the signaled NAL unit size. The NAL unit stream format can be extracted from the sample stream format by traversing through the sample stream format, reading the size information and appropriately extracting each NAL unit. Methods of framing NAL units in a manner other than the use of the sample stream format are outside the scope of this document. The sample stream format is specified in Annex C.

[Ed. (All): Does the current syntax produce V-PCC units in decoding order?]

[Ed. (All): The above question will be clearly addressed in section 8]

## Partitioning of atlas frames, tiles, and tile groups

### Partitioning of atlas frames into tiles and tile groups

This subclause specifies how an atlas frame is partitioned into tiles and tiles groups.

An atlas frame is divided into one or more tile rows and one or more tile columns. A tile is a rectangular region of an atlas frame.

A tile group contains a number of tiles of an atlas frame.

Only rectangular tile groups are supported. In this mode, a tile group contains a number of tiles of an atlas frame that collectively form a rectangular region of the atlas frame.

Figure 6‑1 shows an example tile group partitioning of an atlas frame, where the atlas frame is divided into 24 tiles (6 tile columns and 4 tile rows) and 9 rectangular tile groups.

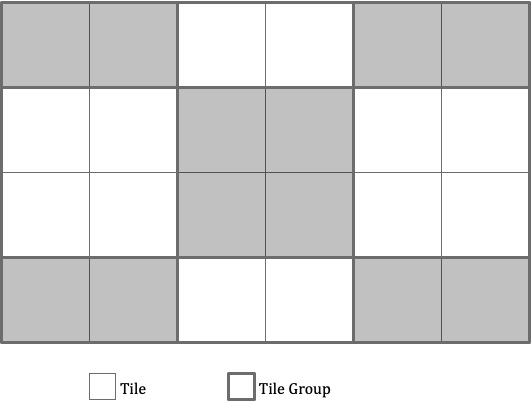


Figure 6‑1 – An atlas frame that is partitioned into 24 tiles and 9 tile groups (informative)

## Scanning Processes

### Tile scanning process

The list ColWidth[ i ] for i ranging from 0 to afti\_num\_tile\_columns\_minus1, inclusive, specifying the width of the i-th tile column in units of 64 samples, is derived, and, when afti\_uniform\_tile\_spacing\_flag is equal to 1, the value of afti\_num\_tile\_columns\_minus1 is inferred, as follows:

if( afti\_uniform\_tile\_spacing\_flag ) {  
 remainingWidthInBlocks = ( asps\_frame\_width + 63 ) / 64  
 i = 0  
 while( remainingWidthInBlocks > ( afti\_tile\_cols\_width\_minus1 + 1 ) ) {  
 ColWidth[ i++ ] = afti\_tile\_cols\_width\_minus1 + 1  
 remainingWidthInBlocks −= ( afti\_tile\_cols\_width\_minus1 + 1 )  
 }  
 ColWidth[ i ] = remainingWidthInBlocks  
 afti\_num\_tile\_columns\_minus1 = i  
} else {  
 ColWidth[ afti\_num\_tile\_columns\_minus1 ] = ( asps\_frame\_width + 63 ) / 64 (6‑1)  
 for( i = 0; i < afti\_num\_tile\_columns\_minus1; i++ ) {  
 ColWidth[ i ] = afti\_tile\_column\_width\_minus1[ i ] + 1  
 ColWidth[ afti\_num\_tile\_columns\_minus1 ] −= ColWidth[ i ]  
 }  
}

The list RowHeight[ j ] for j ranging from 0 to afti\_num\_tile\_rows\_minus1, inclusive, specifying the height of the j-th tile row in units of 64 samples, is derived, and, when afti\_uniform\_tile\_spacing\_flag is equal to 1, the value of afti\_num\_tile\_rows\_minus1 is inferred, as follows:

if( afti\_uniform\_tile\_spacing\_flag ) {  
 remainingHeightInBlocks = ( asps\_frame\_height + 63 ) / 64  
 i = 0  
 while( remainingHeightInBlocks > ( afti\_tile\_rows\_height\_minus1 + 1 ) ) {  
 RowHeight[ i++ ] = afti\_tile\_rows\_height\_minus1 + 1  
 remainingHeightInBlocks −= ( afti\_tile\_rows\_height\_minus1 + 1 )  
 }  
 RowHeight[ i ] = remainingHeightInBlocks  
 afti\_num\_tile\_rows\_minus1 = i  
 } else {  
 RowHeight[ afti\_num\_tile\_rows\_minus1 ] = ( asps\_frame\_height + 63 ) / 64 (6‑2)  
 for( j = 0; j < afti\_num\_tile\_rows\_minus1; j++ ) {  
 RowHeight[ j ] = afti\_tile\_row\_height\_minus1[ j ] + 1  
 RowHeight[ afti\_num\_tile\_rows\_minus1 ] −= RowHeight[ j ]  
 }  
 }

The list TileColBd[ i ] for i ranging from 0 to afti\_num\_tile\_columns\_minus1 + 1, inclusive, specifying the location of the i-th tile column boundary in units of 64 samples, is derived as follows:

for( TileColBd[ 0 ] = 0, i = 0; i <= afti\_num\_tile\_columns\_minus1; i++ )  
 TileColBd[ i + 1 ] = TileColBd[ i ] + ColWidth[ i ] (6‑3)

The list TileRowBd[ j ] for j ranging from 0 to afti\_num\_tile\_rows\_minus1 + 1, inclusive, specifying the location of the j-th tile row boundary in units of 64 samples, is derived as follows:

for( TileRowBd[ 0 ] = 0, j = 0; j <= afti\_num\_tile\_rows\_minus1; j++ )  
 TileRowBd[ j + 1 ] = TileRowBd[ j ] + RowHeight[ j ] (6‑4)

[Ed. (ALL): Add information about the 3D coordinate system that we are currently using.]

[Ed. (ALL): We seem to need an additional subsection in section 6 that would allow us to describe the source, decoded, and output point cloud/atlas frame formats. This could include some high level description of how atlas, occupancy, geometry, and attributes, as well as maps relate.]

[Ed. (ALL): A section is also needed here that would define patches]

# Syntax and semantics

## Method of specifying syntax in tabular form

The syntax tables specify a superset of the syntax of all allowed bitstreams. Additional constraints on the syntax may be specified, either directly or indirectly, in other clauses.

NOTE – An actual decoder should implement some means for identifying entry points into the bitstream and some means to identify and handle non-conforming bitstreams. The methods for identifying and handling errors and other such situations are not specified in this document.

The following table lists examples of the syntax specification format. When **syntax\_element** appears, it specifies that a syntax element is parsed from the bitstream and the bitstream pointer is advanced to the next position beyond the syntax element in the bitstream parsing process.

|  |  |
| --- | --- |
|  | Descriptor |
| /\* A statement can be a syntax element with an associated descriptor or can be an expression used to specify conditions for the existence, type, and quantity of syntax elements, as in the following two examples \*/ |  |
| **syntax\_element** | ue(v) |
| conditioning statement |  |
|  |  |
| /\* A group of statements enclosed in curly brackets is a compound statement and is treated functionally as a single statement. \*/ |  |
| { |  |
| statement |  |
| statement |  |
| ... |  |
| } |  |
|  |  |
| /\* A "while" structure specifies a test of whether a condition is true, and if true, specifies evaluation of a statement (or compound statement) repeatedly until the condition is no longer true \*/ |  |
| while( condition ) |  |
| statement |  |
|  |  |
| /\* A "do ... while" structure specifies evaluation of a statement once, followed by a test of whether a condition is true, and if true, specifies repeated evaluation of the statement until the condition is no longer true \*/ |  |
| do |  |
| statement |  |
| while( condition ) |  |
|  |  |
| /\* An "if ... else" structure specifies a test of whether a condition is true and, if the condition is true, specifies evaluation of a primary statement, otherwise, specifies evaluation of an alternative statement. The "else" part of the structure and the associated alternative statement is omitted if no alternative statement evaluation is needed \*/ |  |
| if( condition ) |  |
| primary statement |  |
| else |  |
| alternative statement |  |
|  |  |
| /\* A "for" structure specifies evaluation of an initial statement, followed by a test of a condition, and if the condition is true, specifies repeated evaluation of a primary statement followed by a subsequent statement until the condition is no longer true. \*/ |  |
| for( initial statement; condition; subsequent statement ) |  |
| primary statement |  |

## Specification of syntax functions and descriptors

The functions presented here are used in the syntactical description. These functions are expressed in terms of the value of a bitstream pointer that indicates the position of the next bit to be read by the decoding process from the bitstream.

byte\_aligned( ) is specified as follows:

– If the current position in the bitstream is on a byte boundary, i.e. the next bit in the bitstream is the first bit in a byte, the return value of byte\_aligned( ) is equal to TRUE.

– Otherwise, the return value of byte\_aligned( ) is equal to FALSE.

more\_data\_in\_payload( ) is specified as follows:

* If byte\_aligned( ) is equal to TRUE and the current position in the sei\_payload( ) syntax structure is 8 \* payloadSize bits from the beginning of the sei\_payload( ) syntax structure, the return value of more\_data\_in\_payload( ) is equal to FALSE.
* Otherwise, the return value of more\_data\_in\_payload( ) is equal to TRUE.

more\_rbsp\_data( ) is specified as follows:

* If there is no more data in the raw byte sequence payload (RBSP), the return value of more\_rbsp\_data( ) is equal to FALSE.
* Otherwise, the RBSP data are searched for the last (least significant, right-most) bit equal to 1 that is present in the RBSP. Given the position of this bit, which is the first bit (rbsp\_stop\_one\_bit) of the rbsp\_trailing\_bits( ) syntax structure, the following applies:
* If there is more data in an RBSP before the rbsp\_trailing\_bits( ) syntax structure, the return value of more\_rbsp\_data( ) is equal to TRUE.
* Otherwise, the return value of more\_rbsp\_data( ) is equal to FALSE.

The method for enabling determination of whether there is more data in the RBSP is specified by the application (or in Annex C for applications that use the sample stream format).

more\_rbsp\_trailing\_data( ) is specified as follows:

* If there is more data in an RBSP, the return value of more\_rbsp\_trailing\_data( ) is equal to TRUE.
* Otherwise, the return value of more\_rbsp\_trailing\_data( ) is equal to FALSE.

more\_data\_in\_vpcc\_unit( ) is specified as follows:

* If more data follow in the current vpcc\_unit, i.e. the decoded data up to now in the current vpcc\_unit is less than numBytesInVPCCInit, the return value of more\_data\_in\_vpcc\_unit( ) is equal to TRUE.
* Otherwise, the return value of more\_data\_in\_vpcc\_unit( ) is equal to FALSE.

next\_bits( n ) provides the next bits in the bitstream for comparison purposes, without advancing the bitstream pointer. Provides a look at the next n bits in the bitstream with n being its argument.

payload\_extension\_present( ) is specified as follows:

* If the current position in the sei\_payload( ) syntax structure is not the position of the last (least significant, right-most) bit that is equal to 1 that is less than 8 \* payloadSize bits from the beginning of the syntax structure (i.e., the position of the payload\_bit\_equal\_to\_one syntax element), the return value of payload\_extension\_present( ) is equal to TRUE.
* Otherwise, the return value of payload\_extension\_present( ) is equal to FALSE.

afrm\_layer\_id( frameX ) returns the value of the nal\_layer\_id of the ACL NAL units in the atlas frame frameX.

read\_bits( n ) reads the next n bits from the bitstream and advances the bitstream pointer by n bit positions. When n is equal to 0, read\_bits( n ) is specified to return a value equal to 0 and to not advance the bitstream pointer.

The following descriptors specify the parsing process of each syntax element:

– b(8): byte having any pattern of bit string (8 bits). The parsing process for this descriptor is specified by the return value of the function read\_bits( 8 ).

– f(n): fixed-pattern bit string using n bits written (from left to right) with the left bit first. The parsing process for this descriptor is specified by the return value of the function read\_bits( n ).

– fl(n): binary floating point value using n bits. The parsing process for this descriptor is according to [IEEE754].

[Ed. Note: Add normative reference to IEEE754]

– i(n): signed integer using n bits. When n is "v" in the syntax table, the number of bits varies in a manner dependent on the value of other syntax elements. The parsing process for this descriptor is specified by the return value of the function read\_bits( n ) interpreted as a two's complement integer representation with most significant bit written first. In particular, the parsing process for this descriptor is specified as follows:

i(n) {   
 value = read\_bits( n )   
 if( value < ( 1 << ( n – 1 )))   
 return value   
 else  
 return ( value | ~(( 1 << (n – 1)) – 1 )   
 }

– se(v): signed integer 0-th order Exp-Golomb-coded syntax element with the left bit first. The parsing process for this descriptor is specified in clause 10.2.

– st(v): null-terminated string encoded as UTF-8 characters as specified in ISO/IEC 10646. The parsing process is specified as follows: st(v) begins at a byte-aligned position in the bitstream and reads and returns a series of bytes from the bitstream, beginning at the current position and continuing up to but not including the next byte-aligned byte that is equal to 0x00, and advances the bitstream pointer by ( stringLength + 1 ) \* 8 bit positions, where stringLength is equal to the number of bytes returned.

NOTE – The st(v) syntax descriptor is only used in this document when the current position in the bitstream is a byte-aligned position.

– u(n): unsigned integer using n bits. When n is "v" in the syntax table, the number of bits varies in a manner dependent on the value of other syntax elements. The parsing process for this descriptor is specified by the return value of the function read\_bits( n ) interpreted as a binary representation of an unsigned integer with the most significant bit written first.

– ue(v): unsigned integer 0-th order Exp-Golomb-coded syntax element with the left bit first. The parsing process for this descriptor is specified in clause 10.2.

## Syntax in tabular form

### General

The V‑PCC bitstream is composed of a set of V-PCC units as shown in Figure 7‑1. Each V-PCC unit has a V‑PCC unit header and a V‑PCC unit payload. The V‑PCC unit header describes the V-PCC unit type. Table 7‑1 describes the supported V‑PCC unit types.

The Attribute Video Data V‑PCC unit header specifies also the attribute type and its index, allowing multiple instances of the same attribute type to be supported. Table 7‑2 specifies the supported V‑PCC attribute types.

The occupancy, geometry, and attribute Video Data unit payloads correspond to video data units (e.g., HEVC NAL unit) that could be decoded by an appropriate video decoder.

[Ed. (AMT): Need to clarify what is an appropriate video decoder. Specifying the decoder may be a process done outside V-PCC specification, e.g. at the system level.]



Figure 7‑1 V-PCC Bitstream structure

[Ed. (AMT): Need to revise figure;]

[Ed. (SS): Describe colour coding of different boxes][Ed. (AMT): Why is that needed? It is just for visualization purposes and easier reading. Indeed some are similar types of data, e.g. video data, but I think it is not needed to describe things further]

[Ed. (SS): Duplicated “sequence parameter set”, rename or motivate duplication][It is atlas sps. Video data also have theor own sequence parameter sets. It should be hopefully obvious that we are referring to that level]

[Ed: (ATA): isn’t it better to rename Geometry/attribute Patch Parameter Set to Patch Geometry/Attribute Parameter Set for name consistency with Frame Geometry/Attribute Parameter Set?]

### V-PCC unit syntax

#### General V-PCC unit syntax

|  |  |
| --- | --- |
| vpcc\_unit( numBytesInVPCCUnit) { | **Descriptor** |
| vpcc\_unit\_header( ) |  |
| vpcc\_unit\_payload( ) |  |
| while( more\_data\_in\_vpcc\_unit ) |  |
| **trailing\_zero\_8bits** /\* equal to 0x00 \*/ | f(8) |
| } |  |

#### V-PCC unit header syntax

|  |  |
| --- | --- |
| vpcc\_unit\_header( ) { | **Descriptor** |
| **vuh\_unit\_type** | u(5) |
| if( vuh\_unit\_type  = =  VPCC\_AVD  | |  vuh\_unit\_type  = =  VPCC\_GVD  | |  vuh\_unit\_type  = =  VPCC\_OVD  | |  vuh\_unit\_type  = =  VPCC\_AD ) { |  |
| **vuh\_vpcc\_parameter\_set\_id** | u(4) |
| **vuh\_atlas\_id** | u(6) |
| } |  |
| if( vuh\_unit\_type  = =  VPCC\_AVD ) { |  |
| **vuh\_attribute\_index** | u(7) |
| **vuh\_attribute\_partition\_index** | u(5) |
| **vuh\_map\_index** | u(4) |
| **vuh\_auxiliary\_video\_flag** | u(1) |
| } else if( vuh\_unit\_type  = =  VPCC\_GVD ) { |  |
| **vuh\_map\_index** | u(4) |
| **vuh\_auxiliary\_video\_flag** | u(1) |
| **vuh\_reserved\_zero\_12bits** | u(12) |
| } else if( vuh\_unit\_type  = =  VPCC\_OVD  | |  vuh\_unit\_type  = =  VPCC\_AD ) |  |
| **vuh\_reserved\_zero\_17bits** | u(17) |
| else |  |
| **vuh\_reserved\_zero\_27bits** | u(27) |
| } |  |

[Ed. (All): It was suggested during the editors meeting that maybe a one bit flag, in the case of vuh\_unit\_type == VPCC\_AD, could be used as an indicator of the V-PCC access unit. This is proposed as an alternative to what is currently being done using a NAL unit type. This may make it easier to parse bitstreams and acknowledge the presence of such access units.]

#### V-PCC unit payload syntax

|  |  |
| --- | --- |
| vpcc\_unit\_payload( ) { | **Descriptor** |
| if( vuh\_unit\_type  = =  VPCC\_VPS ) |  |
| vpcc\_parameter\_set( ) |  |
| else if( vuh\_unit\_type  = =  VPCC\_AD ) |  |
| atlas\_sub\_bitstream( ) |  |
| else if( vuh\_unit\_type  = =  VPCC\_OVD  | |  vuh\_unit\_type  = =  VPCC\_GVD  | |  vuh\_unit\_type  = =  VPCC\_AVD) |  |
| video\_sub\_bitstream( ) |  |
| } |  |

### Byte alignment syntax

|  |  |
| --- | --- |
| byte\_alignment( ) { | **Descriptor** |
| **alignment\_bit\_equal\_to\_one** /\* equal to 1 \*/ | f(1) |
| while( !byte\_aligned( ) ) |  |
| **alignment\_bit\_equal\_to\_zero** /\* equal to 0 \*/ | f(1) |
| } |  |

### Sequence parameter set syntax

#### General V-PCC parameter set syntax

|  |  |
| --- | --- |
| vpcc\_parameter\_set( ) { | **Descriptor** |
| profile\_tier\_level() |  |
| **vps\_vpcc\_parameter\_set\_id** | u(4) |
| **vps\_reserved\_zero\_8bits** | u(8) |
| **vps\_atlas\_count\_minus1** | u(6) |
| for(j = 0; j < vps\_atlas\_count\_minus1 + 1; j++ ) { |  |
| **vps\_frame\_width**[ j ] | u(16) |
| **vps\_frame\_height**[ j ] | u(16) |
| **vps\_map\_count\_minus1**[ j ] | u(4) |
| if( vps\_map\_count\_minus1[ j ] > 0 ) |  |
| **vps\_multiple\_map\_streams\_present\_flag**[ j ] | u(1) |
| vps\_map\_absolute\_coding\_enabled\_flag[ j ][ 0 ] = 1 |  |
| for( i = 1; i <= vps\_map\_count\_minus1[ j ]; i++ ) { |  |
| if( vps\_multiple\_map\_streams\_present\_flag[ j ] ) |  |
| **vps\_map\_absolute\_coding\_enabled\_flag**[ j ][ i ] | u(1) |
| else |  |
| vps\_map\_absolute\_coding\_enabled\_flag[ j ][ i ] = 1 |  |
| if( vps\_map\_absolute\_coding\_enabled\_flag[ j ][ i ]  = =  0 ) { |  |
| if( i > 0) |  |
| **vps\_map\_predictor\_index\_diff**[ j ][ i ] | ue(v) |
| else |  |
| vps\_map\_predictor\_index\_diff[ j ][ i ] = 0 |  |
| } |  |
| } |  |
| **vps\_auxiliary\_video\_present\_flag**[ j ] | u(1) |
| occupancy\_information( j ) |  |
| geometry\_information( j ) |  |
| attribute\_information( j ) |  |
| } |  |
| **vps\_extension\_present\_flag** | u(1) |
| if(vps\_extension\_present\_flag) { |  |
| **vps\_extension\_length\_minus1** | ue(v) |
| for( j = 0; j < vps\_extension\_length\_minus1 + 1; j++ ) { |  |
| **vps\_extension\_data\_byte** | u(8) |
| } |  |
| } |  |
| byte\_alignment( ) |  |
| } |  |

[Ed. (AMT): Need to update frame rate parameters using a concept similar to vps\_num\_units\_in\_tick in AVC/HEVC. Frame rate should also be included in the ASPS.]

#### Profile, tier, and level syntax

|  |  |
| --- | --- |
| profile\_tier\_level( ) { | **Descriptor** |
| **ptl\_tier\_flag** | u(1) |
| **ptl\_profile\_codec\_group\_idc** | u(7) |
| **ptl\_profile\_pcc\_toolset\_idc** | u(8) |
| **ptl\_profile\_reconstruction\_idc** | u(8) |
| **ptl\_reserved\_zero\_32bits** | u(32) |
| **ptl\_level\_idc** | u(8) |
| } |  |

#### Occupancy information syntax

|  |  |
| --- | --- |
| occupancy\_information( atlasId ) { | **Descriptor** |
| **oi\_occupancy\_codec\_id**[ atlasId ] | u(8) |
| **oi\_lossy\_occupancy\_map\_compression\_threshold**[ atlasId ] | u(8) |
| **oi\_occupancy\_nominal\_2d\_bitdepth\_minus1**[ atlasId ] | u(5) |
| **oi\_occupancy\_MSB\_align\_flag**[ atlasId ] | u(1) |
| } |  |

#### Geometry information syntax

|  |  |
| --- | --- |
| geometry\_information(atlasId ) { | **Descriptor** |
| **gi\_geometry\_codec\_id**[ atlasId ] | u(8) |
| **gi\_geometry\_nominal\_2d\_bitdepth\_minus1**[ atlasId ] | u(5) |
| **gi\_geometry\_MSB\_align\_flag**[ atlasId ] | u(1) |
| **gi\_geometry\_3d\_coordinates\_bitdepth\_minus1**[ atlasId ] | u(5) |
| if( vps\_auxiliary\_video\_present\_flag[ atlasId ] ) |  |
| **gi\_auxiliary\_geometry\_codec\_id**[ atlasId ] | u(8) |
| } |  |

#### Attribute information syntax

|  |  |
| --- | --- |
| attribute\_information(atlasId) { | **Descriptor** |
| **ai\_attribute\_count**[ atlasId ] | u(7) |
| for( i = 0; i < ai\_attribute\_count[ atlasId ]; i++ ) { |  |
| **ai\_attribute\_type\_id**[ atlasId ][ i ] | u(4) |
| **ai\_attribute\_codec\_id**[ atlasId ][ i ] | u(8) |
| if( vps\_auxiliary\_video\_present\_flag[ atlasId ] ) |  |
| **ai\_auxiliary\_attribute\_codec\_id**[ atlasId ][ i ] | u(8) |
| if( vps\_map\_count\_minus1[ atlasId ] > 0 ) |  |
| **ai\_attribute\_map\_absolute\_coding\_persistence\_flag**[ atlasId ][ i ] | u(1) |
| **ai\_attribute\_dimension\_minus1**[ atlasId ][ i ] | u(6) |
| if( ai\_attribute\_dimension\_minus1[ atlasId ][ i ] > 0 ) { |  |
| **ai\_attribute\_dimension\_partitions\_minus1**[ atlasId ][ i ] | u(6) |
| remainingDimensions = ai\_attribute\_dimension\_minus1[ atlasId ][ i ] |  |
| k = ai\_attribute\_dimension\_partitions\_minus1[ atlasId ][ i ] |  |
| for( j = 0; j < k; j++ ) { |  |
| if( k – j == remainingDimensions ) |  |
| ai\_attribute\_partition\_channels\_minus1[ atlasId ][ i ][ j ] = 0 |  |
| else |  |
| **ai\_attribute\_partition\_channels\_minus1**[ atlasId ][ i ][ j ] | ue(v) |
| remainingDimensions –=   ai\_attribute\_partition\_channels\_minus1[ atlasId ][ i ][ j ] + 1 |  |
| } |  |
| ai\_attribute\_partition\_channels\_minus1[ atlasId ][ i ][ k ] = remainingDimensions |  |
| } |  |
| **ai\_attribute\_nominal\_2d\_bitdepth\_minus1**[ atlasId ][ i ] | u(5) |
| **ai\_attribute\_MSB\_align\_flag**[ atlasId ][ i ] | u(1) |
| } |  |
| } |  |

[Ed. (RJ): Flags similar to ai\_attribute\_MSB\_align\_flag should be considered for occupancy map and RAW video. Furthermore, consider signalling nominal\_2d\_bitdepth\_minus1 for occupancy map as well as RAW video.][Ed. (RJ): MSB\_align\_flag and nominal\_2d\_bitdepth\_flag have been introduced for geometry as well as occupancy map. There are no such flags for auxiliary video. Perhaps they should be considered?]

### NAL unit syntax

#### General NAL unit syntax

|  |  |
| --- | --- |
| nal\_unit( NumBytesInNalUnit ) { | **Descriptor** |
| nal\_unit\_header( ) |  |
| NumBytesInRbsp = 0 |  |
| for( i = 2; i < NumBytesInNalUnit; i++ ) |  |
| **rbsp\_byte**[ NumBytesInRbsp++ ] | b(8) |
| } |  |

#### NAL unit header syntax

|  |  |
| --- | --- |
| nal\_unit\_header( ) { | **Descriptor** |
| **nal\_forbidden\_zero\_bit** | f(1) |
| **nal\_unit\_type** | u(6) |
| **nal\_layer\_id** | u(6) |
| **nal\_temporal\_id\_plus1** | u(3) |
| } |  |

### Raw byte sequence payloads, trailing bits, and byte alignment syntax

#### Atlas sequence parameter set RBSP syntax

##### General atlas sequence parameter set RBSP syntax

|  |  |
| --- | --- |
| atlas\_sequence\_parameter\_set\_rbsp( ) { | **Descriptor** |
| **asps\_atlas\_sequence\_parameter\_set\_id** | ue(v) |
| **asps\_frame\_width** | u(16) |
| **asps\_frame\_height** | u(16) |
| **asps\_log2\_patch\_packing\_block\_size** | u(3) |
| **asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4** | ue(v) |
| **asps\_max\_dec\_atlas\_frame\_buffering\_minus1** | ue(v) |
| **asps\_long\_term\_ref\_atlas\_frames\_flag** | u(1) |
| **asps\_num\_ref\_atlas\_frame\_lists\_in\_asps** | ue(v) |
| for( i = 0; i < asps\_num\_ref\_atlas\_frame\_lists\_in\_asps; i++ ) |  |
| ref\_list\_struct( i ) |  |
| **asps\_use\_eight\_orientations\_flag** | u(1) |
| **asps\_extended\_projection\_enabled\_flag** | u(1) |
| **asps\_normal\_axis\_limits\_quantization\_enabled\_flag** | u(1) |
| **asps\_normal\_axis\_max\_delta\_value\_enabled\_flag** | u(1) |
| **asps\_remove\_duplicate\_point\_enabled\_flag** | u(1) |
| **asps\_pixel\_deinterleaving\_enabled\_flag** | u(1) |
| **asps\_patch\_precedence\_order\_flag** | u(1) |
| **asps\_patch\_size\_quantizer\_present\_flag** | u(1) |
| **asps\_raw\_patch\_enabled\_flag** | u(1) |
| **asps\_eom\_patch\_enabled\_flag** | u(1) |
| if (asps\_raw\_patch\_enabled\_flag || asps\_eom\_patch\_enabled\_flag ) |  |
| **asps\_auxiliary\_video\_enabled\_flag** | u(1) |
| **asps\_point\_local\_reconstruction\_enabled\_flag** | u(1) |
| **asps\_map\_count\_minus1** | u(4) |
| if( asps\_pixel\_deinterleaving\_enabled\_flag ) |  |
| for( j = 0; j < = asps\_map\_count\_minus1; j++ ) |  |
| **asps\_pixel\_deinterleaving\_map\_flag**[ j ] | u(1) |
| if( asps\_eom\_patch\_enabled\_flag  && asps\_map\_count\_minus1 = = 0 ) |  |
| **asps\_eom\_fix\_bit\_count\_minus1** | u(4) |
| if( asps\_point\_local\_reconstruction\_enabled\_flag ) |  |
| asps\_point\_local\_reconstruction\_information( asps\_map\_count\_minus1 ) |  |
| if( asps\_pixel\_deinterleaving\_enabled\_flag || asps\_point\_local\_reconstruction\_enabled\_flag ) |  |
| **asps\_surface\_thickness\_minus1** | u(8) |
| **asps\_vui\_parameters\_present\_flag** | u(1) |
| if( asps\_vui\_parameters\_present\_flag ) |  |
| vui\_parameters( ) |  |
| **asps\_extension\_flag** | u(1) |
| if( asps\_extension\_flag ) |  |
| while( more\_rbsp\_data( ) ) |  |
| **asps\_extension\_data\_flag** | u(1) |
| rbsp\_trailing\_bits( ) |  |
| } |  |

##### Point local reconstruction information syntax

|  |  |
| --- | --- |
| asps\_point\_local\_reconstruction\_information( mapCountMinus1 ) { | **Descriptor** |
| for( i = 0; i < mapCountMinus1 + 1 ; i ++ ) { |  |
| **plri\_point\_local\_reconstruction\_map\_flag**[ i ] | u(1) |
| if( plri\_point\_local\_reconstruction\_map\_flag[ i ] ) { |  |
| **plri\_number\_of\_modes\_minus1**[ i ] | u(4) |
| for( j = 0; j < plri\_number\_of\_modes\_minus1[ i ] + 1; j++ ) { |  |
| **plri\_interpolate\_flag**[ i ][ j ] | u(1) |
| **plri\_filling\_flag**[ i ][ j ] | u(1) |
| **plri\_minimum\_depth**[ i ][ j ] | u(2) |
| **plri\_neighbour\_minus1**[ i ][ j ] | u(2) |
| } |  |
| **plri\_block\_threshold\_per\_patch\_minus1**[ mapIdx ] | u(6) |
| } |  |
| } |  |
| } |  |

#### Atlas frame parameter set RBSP syntax

##### General atlas frame parameter set RBSP syntax

|  |  |
| --- | --- |
| atlas\_frame\_parameter\_set\_rbsp( ) { | **Descriptor** |
| **afps\_atlas\_frame\_parameter\_set\_id** | ue(v) |
| **afps\_atlas\_sequence\_parameter\_set\_id** | ue(v) |
| atlas\_frame\_tile\_information( ) |  |
| **afps\_output\_flag\_present\_flag** | u(1) |
| **afps\_num\_ref\_idx\_default\_active\_minus1** | ue(v) |
| **afps\_additional\_lt\_afoc\_lsb\_len** | ue(v) |
| **afps\_3d\_pos\_x\_bit\_count\_minus1** | u(5) |
| **afps\_3d\_pos\_y\_bit\_count\_minus1** | u(5) |
| **afps\_lod\_mode\_enabled\_flag** | u(1) |
| **afps\_override\_eom\_for\_depth\_flag** | u(1) |
| if( afps\_override\_eom\_for\_depth\_flag ) { |  |
| **afps\_eom\_number\_of\_patch\_bit\_count\_minus1** | u(4) |
| **afps\_eom\_max\_bit\_count\_minus1** | u(4) |
| } |  |
| **afps\_raw\_3d\_pos\_bit\_count\_explicit\_mode\_flag** | u(1) |
| **afps\_extension\_flag** | u(1) |
| if( afps\_extension\_flag ) |  |
| while( more\_rbsp\_data( ) ) |  |
| **afps\_extension\_data\_flag** | u(1) |
| rbsp\_trailing\_bits( ) |  |
| } |  |

##### Atlas frame tile information syntax

|  |  |
| --- | --- |
| atlas\_frame\_tile\_information( ) { | **Descriptor** |
| **afti\_single\_tile\_in\_atlas\_frame\_flag** | u(1) |
| if( !afti\_single\_tile\_in\_atlas\_frame\_flag ) { |  |
| **afti\_uniform\_tile\_spacing\_flag** | u(1) |
| if( afti\_uniform\_tile\_spacing\_flag ) { |  |
| **afti\_tile\_cols\_width\_minus1** | ue(v) |
| **afti\_tile\_rows\_height\_minus1** | ue(v) |
| } else { |  |
| **afti\_num\_tile\_columns\_minus1** | ue(v) |
| **afti\_num\_tile\_rows\_minus1** | ue(v) |
| for( i = 0; i < afti\_num\_tile\_columns\_minus1; i++ ) |  |
| **afti\_tile\_column\_width\_minus1**[ i ] | ue(v) |
| for( i = 0; i < afti\_num\_tile\_rows\_minus1; i++ ) |  |
| **afti\_tile\_row\_height\_minus1**[ i ] | ue(v) |
| } |  |
| **afti\_single\_tile\_per\_tile\_group\_flag** | u(1) |
| if( !afti\_single\_tile\_per\_tile\_group\_flag ) { |  |
| **afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1** | ue(v) |
| for( i = 0; i < afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1 + 1; i++ ) { |  |
| if( i > 0 ) |  |
| **afti\_top\_left\_tile\_idx**[ i ] | u(v) |
| **afti\_bottom\_right\_tile\_idx\_delta**[ i ] | u(v) |
| } |  |
| } |  |
| **afti\_signalled\_tile\_group\_id\_flag** | u(1) |
| if( afti\_signalled\_tile\_group\_id\_flag ) { |  |
| **afti\_signalled\_tile\_group\_id\_length\_minus1** | ue(v) |
| for( i = 0; i < afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1 + 1; i++ ) |  |
| **afti\_tile\_group\_id**[ i ] | u(v) |
| } |  |
| } |  |
| } |  |

#### Atlas adaptation parameter set RBSP syntax

##### General atlas adaptation parameter set RBSP syntax

|  |  |
| --- | --- |
| atlas\_adaptation\_parameter\_set\_rbsp() { | **Descriptor** |
| **aaps\_atlas\_adaptation\_parameter\_set\_id** | ue(v) |
| **~~aaps\_atlas\_sequence\_parameter\_set\_id~~** | ~~ue(v)~~ |
| **aaps\_camera\_parameters\_present\_flag** | u(1) |
| if ( aaps\_camera\_parameters\_present\_flag ) |  |
| atlas\_camera\_parameters( ) |  |
| **aaps\_extension\_flag** | u(1) |
| if( aaps\_extension\_flag ) |  |
| while( more\_rbsp\_data( ) ) |  |
| **aaps\_extension\_data\_flag** | u(1) |
| rbsp\_trailing\_bits( ) |  |
| } |  |

##### Atlas camera parameters syntax

|  |  |
| --- | --- |
| atlas\_camera\_parameters() { | **Descriptor** |
| **acp\_camera\_model** | u(8) |
| if( acp\_camera\_model == 1) { |  |
| **acp\_scale\_enabled\_flag** | u(1) |
| **acp\_offset\_enabled\_flag** | u(1) |
| **acp\_rotation\_enabled\_flag** | u(1) |
| if( acp\_scale\_enabled\_flag ) |  |
| for( d = 0; d < 3; d++ ) |  |
| **acp\_scale\_on\_axis**[ d ] | u(32) |
| if( acp\_offset\_enabled\_flag ) |  |
| for( d = 0; d < 3; d++ ) |  |
| **acp\_offset\_on\_axis**[ d ] | i(32) |
| if( acp\_rotation\_enabled\_flag ) { |  |
| **acp\_rotation\_qx** | i(16) |
| **acp\_rotation\_qy** | i(16) |
| **acp\_rotation\_qz** | i(16) |
| } |  |
| } |  |
| } |  |

#### Supplemental enhancement information RBSP syntax

|  |  |
| --- | --- |
| sei\_rbsp( ) { | **Descriptor** |
| do |  |
| sei\_message( ) |  |
| while( more\_rbsp\_data( ) ) |  |
| rbsp\_trailing\_bits( ) |  |
| } |  |

#### Access unit delimiter RBSP syntax

|  |  |
| --- | --- |
| access\_unit\_delimiter\_rbsp( ) { | **Descriptor** |
| **aframe\_type** | u(3) |
| rbsp\_trailing\_bits( ) |  |
| } |  |

#### End of sequence RBSP syntax

|  |  |
| --- | --- |
| end\_of\_sequence\_rbsp( ) { | **Descriptor** |
| } |  |

#### End of bitstream RBSP syntax

|  |  |
| --- | --- |
| end\_of\_atlas\_sub\_bitstream\_rbsp( ) { | **Descriptor** |
| } |  |

#### Filler data RBSP syntax

|  |  |
| --- | --- |
| filler\_data\_rbsp( ) { | **Descriptor** |
| while( next\_bits( 8 ) = = 0xFF ) |  |
| **ff\_byte** /\* equal to 0xFF \*/ | f(8) |
| rbsp\_trailing\_bits( ) |  |
| } |  |

#### Atlas tile group layer RBSP syntax

|  |  |
| --- | --- |
| atlas\_tile\_group\_layer\_rbsp( ) { | Descriptor |
| atlas\_tile\_group\_header( ) |  |
| if( atgh\_type != SKIP\_TILE\_GRP ) |  |
| atlas\_tile\_group\_data\_unit( ) |  |
| rbsp\_trailing\_bits( ) |  |
| } |  |

#### Atlas tile group header syntax

|  |  |
| --- | --- |
| atlas\_tile\_group\_header( ) { | **Descriptor** |
| **atgh\_atlas\_frame\_parameter\_set\_id** | ue(v) |
| **atgh\_atlas\_adaptation\_parameter\_set\_id** | ue(v) |
| **atgh\_address** | u(v) |
| **atgh\_type** | ue(v) |
| if( afps\_output\_flag\_present\_flag ) |  |
| **atgh\_atlas\_output\_flag** | u(1) |
| **atgh\_atlas\_frm\_order\_cnt\_lsb** | u(v) |
| if( asps\_num\_ref\_atlas\_frame\_lists\_in\_asps > 0 ) |  |
| **atgh\_ref\_atlas\_frame\_list\_sps\_flag** | u(1) |
| if( atgh\_ref\_atlas\_frame\_list\_sps\_flag == 0) |  |
| ref\_list\_struct( asps\_num\_ref\_atlas\_frame\_lists\_in\_asps ) |  |
| else if( asps\_num\_ref\_atlas\_frame\_lists\_in\_asps > 1 ) |  |
| **atgh\_ref\_atlas\_frame\_list\_idx** | u(v) |
| for( j = 0; j < NumLtrAtlasFrmEntries; j++ ) { |  |
| **atgh\_additional\_afoc\_lsb\_present\_flag**[ j ] | u(1) |
| if( atgh\_additional\_afoc\_lsb\_present\_flag[ j ] ) |  |
| **atgh\_additional\_afoc\_lsb\_val**[ j ] | u(v) |
| } |  |
| if( atgh\_type != SKIP\_TILE\_GRP ) { |  |
| if( asps\_normal\_axis\_limits\_quantization\_enabled\_flag ) { |  |
| **atgh\_pos\_min\_z\_quantizer** | u(5) |
| if( asps\_normal\_axis\_max\_delta\_value\_enabled\_flag ) |  |
| **atgh\_pos\_delta\_max\_z\_quantizer** | u(5) |
| } |  |
| if( asps\_patch\_size\_quantizer\_present\_flag) { |  |
| **atgh\_patch\_size\_x\_info\_quantizer** | u(3) |
| **atgh\_patch\_size\_y\_info\_quantizer** | u(3) |
| } |  |
| if( afps\_raw\_3d\_pos\_bit\_count\_explicit\_mode\_flag ) |  |
| **atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1** | u(v) |
| if( atgh\_type == P\_TILE\_GRP && num\_ref\_entries[ RlsIdx ] > 1 ) { |  |
| **atgh\_num\_ref\_idx\_active\_override\_flag** | u(1) |
| if( atgh\_num\_ref\_idx\_active\_override\_flag ) |  |
| **atgh\_num\_ref\_idx\_active\_minus1** | ue(v) |
| } |  |
| } |  |
| byte\_alignment( ) |  |
| } |  |

#### Reference list structure syntax

|  |  |
| --- | --- |
| ref\_list\_struct( rlsIdx ) { | Descriptor |
| **num\_ref\_entries**[ rlsIdx ] | ue(v) |
| for( i = 0; i < num\_ref\_entries[ rlsIdx ]; i++ ) { |  |
| if( asps\_long\_term\_ref\_atlas\_frames\_flag ) |  |
| **st\_ref\_atlas\_frame\_flag**[ rlsIdx ][ i ] | u(1) |
| if( st\_ref\_atlas\_frame\_flag[ rlsIdx ][ i ] ) { |  |
| **abs\_delta\_afoc\_st**[ rlsIdx ][ i ] | ue(v) |
| if( abs\_delta\_afoc\_st[ rlsIdx ][ i ] > 0 ) |  |
| **strpf\_entry\_sign\_flag**[ rlsIdx ][ i ] | u(1) |
| } else |  |
| **afoc\_lsb\_lt**[ rlsIdx ][ i ] | u(v) |
| } |  |
| } |  |

### Atlas tile group data unit syntax

#### General atlas tile group data unit syntax

|  |  |
| --- | --- |
| atlas\_tile\_group\_data\_unit( ) { | **Descriptor** |
| p = 0 |  |
| **atgdu\_patch\_mode**[ p ] | ue(v) |
| while( atgdu\_patch\_mode[ p ] != I\_END && atgdu\_patch\_mode[ p ] != P\_END ){ |  |
| patch\_information\_data( p, atgdu\_patch\_mode[ p ] ) |  |
| p++ |  |
| **atgdu\_patch\_mode**[ p ] | ue(v) |
| } |  |
| AtgduTotalNumberOfPatches = p |  |
| byte\_alignment( ) |  |
| } |  |

#### Patch information data syntax

|  |  |
| --- | --- |
| patch\_information\_data ( patchIdx, patchMode ) { | **Descriptor** |
| if( atgh\_type == SKIP\_TILE\_GR ) |  |
| skip\_patch\_data\_unit( patchIdx ) |  |
| else if( atgh\_type == P\_TILE\_GR ) { |  |
| if( patchMode == P\_SKIP ) |  |
| skip\_patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == P\_MERGE ) |  |
| merge\_patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == P\_INTRA ) |  |
| patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == P\_INTER ) |  |
| inter\_patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == P\_RAW ) |  |
| raw\_patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == P\_EOM ) |  |
| eom\_patch\_data\_unit( patchIdx ) |  |
| } |  |
| else if( atgh\_type == I\_TILE\_GR ) { |  |
| if( patchMode == I\_INTRA ) |  |
| patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == I\_RAW ) |  |
| raw\_patch\_data\_unit( patchIdx ) |  |
| else if( patchMode == I\_EOM ) |  |
| eom\_patch\_data\_unit( patchIdx ) |  |
| } |  |
| } |  |

#### Patch data unit syntax

|  |  |
| --- | --- |
| patch\_data\_unit( patchIdx ) { | **Descriptor** |
| **pdu\_2d\_pos\_x**[ patchIdx ] | ue(v) |
| **pdu\_2d\_pos\_y**[ patchIdx ] | ue(v) |
| **pdu\_2d\_size\_x\_minus1**[ patchIdx ] | ue(v) |
| **pdu\_2d\_size\_y\_minus1**[ patchIdx ] | ue(v) |
| **pdu\_3d\_pos\_x**[ patchIdx ] | u(v) |
| **pdu\_3d\_pos\_y**[ patchIdx ] | u(v) |
| **pdu\_3d\_pos\_min\_z**[ patchIdx ] | u(v) |
| if( asps\_normal\_axis\_max\_delta\_value\_enabled\_flag ) |  |
| **pdu\_3d\_pos\_delta\_max\_z**[ patchIdx ] | u(v) |
| **pdu\_projection\_id**[ patchIdx ] | u(v) |
| **pdu\_orientation\_index**[ patchIdx ] | u(v) |
| if( afps\_lod\_mode\_enabled\_flag ) { |  |
| **pdu\_lod\_enabled\_flag**[ patchIndex ] | u(1) |
| if( pdu\_lod\_enabled\_flag[ patchIndex ] > 0 ) { |  |
| **pdu\_lod\_scale\_x\_minus1**[ patchIndex ] | ue(v) |
| **pdu\_lod\_scale\_y**[ patchIndex ] | ue(v) |
| } |  |
| } | u(v) |
| if( asps\_point\_local\_reconstruction\_enabled\_flag ) |  |
| point\_local\_reconstruction\_data( patchIdx ) |  |
| } |  |

[Ed. (RJ): There seems to be a mismatch between the specification and software for patch orientation.]

#### Skip patch data unit syntax

|  |  |
| --- | --- |
| skip\_patch\_data\_unit( patchIdx ) { | **Descriptor** |
| } |  |

#### Merge patch data unit syntax

|  |  |
| --- | --- |
| merge\_patch\_data\_unit( patchIdx ) { | **Descriptor** |
| if( NumRefIdxActive > 1 ) |  |
| **mpdu\_ref\_index**[ patchIdx ] | ue(v) |
| overridePlrFlag = 0 |  |
| **mpdu\_override\_2d\_params\_flag**[ patchIdx ] | u(1) |
| if( mpdu\_override\_2d\_params\_flag[ patchIdx ] ){ |  |
| **mpdu\_2d\_pos\_x**[ patchIdx ] | se(v) |
| **mpdu\_2d\_pos\_y**[ patchIdx ] | se(v) |
| **mpdu\_2d\_delta\_size\_x**[ patchIdx ] | se(v) |
| **mpdu\_2d\_delta\_size\_y**[ patchIdx ] | se(v) |
| if( asps\_point\_local\_reconstruction\_enabled\_flag ) |  |
| overridePlrFlag = 1 |  |
| } else { |  |
| **mpdu\_override\_3d\_params\_flag**[ patchIdx ] | u(1) |
| if(mpdu\_override\_3d\_params\_flag[ patchIdx ] ) { |  |
| **mpdu\_3d\_pos\_x**[ patchIdx ] | se(v) |
| **mpdu\_3d\_pos\_y**[ patchIdx ] | se(v) |
| **mpdu\_3d\_pos\_min\_z**[ patchIdx ] | se(v) |
| if( asps\_normal\_axis\_max\_delta\_value\_enabled\_flag ) |  |
| **mpdu\_3d\_pos\_delta\_max\_z**[ patchIdx ] | se(v) |
| if( asps\_point\_local\_reconstruction\_enabled\_flag ) { |  |
| **mpdu\_override\_plr\_flag**[ patchIdx ] | u(1) |
| overridePlrFlag = mpdu\_override\_plr\_flag[ patchIdx ] |  |
| } |  |
| } |  |
| } |  |
| if( overridePlrFlag && asps\_point\_local\_reconstruction\_enabled\_flag ) |  |
| point\_local\_reconstruction\_data( patchIdx ) |  |
| } |  |

#### Inter patch data unit syntax

|  |  |
| --- | --- |
| inter\_patch\_data\_unit( patchIdx ) { | **Descriptor** |
| if( NumRefIdxActive > 1 ) |  |
| **ipdu\_ref\_index**[ patchIdx ] | ue(v) |
| **ipdu\_patch\_index**[ patchIdx ] | se(v) |
| **ipdu\_2d\_pos\_x**[ patchIdx  ] | se(v) |
| **ipdu\_2d\_pos\_y**[ patchIdx ] | se(v) |
| **ipdu\_2d\_delta\_size\_x**[ patchIdx ] | se(v) |
| **ipdu\_2d\_delta\_size\_y**[ patchIdx ] | se(v) |
| **ipdu\_3d\_pos\_x**[ patchIdx ] | se(v) |
| **ipdu\_3d\_pos\_y**[ patchIdx ] | se(v) |
| **ipdu\_3d\_pos\_min\_z**[ patchIdx ] | se(v) |
| if( asps\_normal\_axis\_max\_delta\_value\_enabled\_flag ) |  |
| **ipdu\_3d\_pos\_delta\_max\_z**[ patchIdx ] | se(v) |
| if( asps\_point\_local\_reconstruction\_enabled\_flag ) |  |
| point\_local\_reconstruction\_data( patchIdx ) |  |
| } |  |

#### RAW patch data unit syntax

|  |  |
| --- | --- |
| raw\_patch\_data\_unit( patchIdx ) { | **Descriptor** |
| if( asps\_auxiliary\_video\_enabled\_flag ) |  |
| **rpdu\_patch\_in\_auxiliary\_video\_flag**[ patchIdx ] | u(1) |
| **rpdu\_2d\_pos\_x**[ patchIdx ] | ue(v) |
| **rpdu\_2d\_pos\_y**[ patchIdx ] | ue(v) |
| **rpdu\_2d\_size\_x\_minus1**[ patchIdx ] | ue(v) |
| **rpdu\_2d\_size\_y\_minus1**[ patchIdx ] | ue(v) |
| **rpdu\_3d\_pos\_x**[ patchIdx ] | u(v) |
| **rpdu\_3d\_pos\_y**[ patchIdx ] | u(v) |
| **rpdu\_3d\_pos\_z**[ patchIdx ] | u(v) |
| **rpdu\_points\_minus1**[ patchIdx ] | ue(v) |
| } |  |

[Ed. (AMT): Currently the software does not match the specification since it defines two different modes for RAW patches. Determine whether the software should be changed to match the specification or vice versa. ]

[Ed. (AMT): Changed rpdu\_points to ue(v) from se(v) since it does not make sense to be using a negative value. Should we use a fixed representation and define a bit count for it instead? It might save some bits by doing so but caps the maximum number of RAW points that can be signalled based on the fixed representatio case. Btw, can this be zero? If not, should this be rpdu\_raw\_points\_minus1]

[Ed. (AMT): the current implementation does not support creating patches in both the current and in an auxiliar video at the same time. This is required by the decoder and it might be useful in the encoder for testing the support of this functionality.]

#### EOM patch data unit syntax

|  |  |
| --- | --- |
| eom\_patch\_data\_unit( patchIdx ) { | **Descriptor** |
| if( asps\_auxiliary\_video\_enabled\_flag ) |  |
| **epdu\_patch\_in\_auxiliary\_video\_flag**[ patchIdx ] | u(1) |
| **epdu\_2d\_pos\_x**[ patchIdx ] | ue(v) |
| **epdu\_2d\_pos\_y**[ patchIdx ] | ue(v) |
| **epdu\_2d\_size\_x\_minus1**[ patchIdx ] | ue(v) |
| **epdu\_2d\_size\_y\_minus1**[ patchIdx ] | ue(v) |
| **epdu\_patch\_count\_minus1**[ patchIdx ] | u(v) |
| for( i = 0; i < epdu\_patch\_count\_minus1[ patchIdx ] + 1; i++ ) { |  |
| **epdu\_associated\_patch\_idx**[ patchIdx ][ i ] | ue(v) |
| **epdu\_points**[ patchIdx ][ i ] | u(v) |
| } |  |
| } |  |

#### Point local reconstruction data syntax

|  |  |
| --- | --- |
| point\_local\_reconstruction\_data( patchIdx ) { | **Descriptor** |
| for( i = 0; i < asps\_map\_count\_minus1 + 1; i++ ) { |  |
| if( plri\_point\_local\_reconstruction\_map\_flag[ i ] ) { |  |
| if( BlockCount > plri\_block\_threshold\_per\_patch\_minus1[ i ] + 1 ) |  |
| **plrd\_level**[ i ][ patchIdx ] | u(1) |
| else |  |
| plrd\_level[ i ][ patchIdx ] = 1 |  |
| if( plrd\_level[ i ][ patchIdx ] == 0 ) { |  |
| for( j = 0; j < BlockCount; j++ ) { |  |
| **plrd\_present\_block\_flag**[ i ][ patchIdx ][ j ] | u(1) |
| if( plrd\_present\_block\_flag[ i ][ patchIdx ][ j ] ) { |  |
| **plrd\_block\_mode\_minus1**[ i ][ patchIdx ][ j ] | u(v) |
| } |  |
| } |  |
| } else { |  |
| **plrd\_present\_flag**[ i ][ patchIdx ] | u(1) |
| if( plrd\_present\_flag[ i ][ patchIdx ] ) |  |
| **plrd\_mode\_minus1**[ i ][ patchIdx ] | u(v) |
| } |  |
| } |  |
| } |  |
| } |  |

### Supplemental enhancement information message syntax

|  |  |
| --- | --- |
| sei\_message( ) { | **Descriptor** |
| payloadType = 0 |  |
| do { |  |
| **sm\_payload\_type\_byte** | u(8) |
| payloadType += sm\_payload\_type\_byte |  |
| } while( sm\_payload\_type\_byte == 0xFF ) |  |
| payloadSize = 0 |  |
| do{ |  |
| **sm\_payload\_size\_byte** | u(8) |
| payloadSize += sm\_payload\_size\_byte |  |
| } while( sm\_payload\_size\_byte == 0xFF ) |  |
| sei\_payload( payloadType, payloadSize ) |  |
| } |  |

## Semantics

### General

Semantics associated with the syntax structures and with the syntax elements within these structures are specified in this subclause. When the semantics of a syntax element are specified using a table or a set of tables, any values that are not specified in the table(s) shall not be present in the bitstream unless otherwise specified in this document.

### V-PCC unit semantics

#### General V-PCC unit semantics

**trailing\_zero\_8bits** is a byte equal to 0x00.

#### V-PCC unit header semantics

Table 7‑1 V-PCC Unit Types

|  |  |  |  |
| --- | --- | --- | --- |
| **vuh\_unit\_type** | **Identifier** | **V-PCC Unit Type** | **Description** |
| 0 | VPCC\_VPS | V-PCC parameter set | V-PCC level parameters |
| 1 | VPCC\_AD | Atlas data | Atlas information |
| 2 | VPCC\_OVD | Occupancy Video Data | Occupancy information |
| 3 | VPCC\_GVD | Geometry Video Data | Geometry information |
| 4 | VPCC\_AVD | Attribute Video Data | Attribute information |
| 5…31 | VPCC\_RSVD | Reserved | - |

**vuh\_unit\_type** indicates the V-PCC unit type as specified in Table 7‑1.

**vuh\_vpcc\_parameter\_set\_id** specifies the value of vps\_vpcc\_parameter\_set\_id for the active V-PCC VPS. The value of vuh\_vpcc\_parameter\_set\_id shall be in the range of 0 to 15, inclusive.

**vuh\_atlas\_id** specifies the index of the atlas that corresponds to the current V-PCC unit. The value of vuh\_atlas\_id shall be in the range of 0 to 63, inclusive.

**vuh\_attribute\_index** indicates the index of the attribute data carried in the Attribute Video Data unit. The value of vuh\_attribute\_index shall be in the range of 0 to (ai\_attribute\_count[ vuh\_atlas\_id ] – 1), inclusive.

**vuh\_attribute\_partition\_index** indicates the index of the attribute dimension group carried in the Attribute Video Data unit. The value of vuh\_attribute\_partition\_index shall be in the range of 0 to ai\_attribute\_dimension\_partitions\_minus1[ vuh\_atlas\_id ][ vuh\_attribute\_index ], inclusive.

**vuh\_map\_index** when present, indicates the map index of the current geometry or attribute stream. When not present, the map index of the current geometry or attribute stream is derived based on the type of the stream and the operations described in clauses 8.3 and 8.4 for geometry and attribute video streams respectively. The value of vuh\_map\_index, when present, shall be in the range of 0 to vps\_map\_count\_minus1[ vuh\_atlas\_id ], inclusive.

**vuh\_auxiliary\_video\_flag** equal to 1 indicates that the associated geometry or attribute video data unit is a RAW and/or EOM coded points video only. vuh\_auxiliary\_video\_flag equal to 0 indicates that the associated geometry or attribute video data unit may contain RAW and/or EOM coded points. When vuh\_auxiliary\_video\_flag is not present, its value shall be inferred to be equal to 0.

[Ed. (RJ): In semantics, in a number of places, we have the language "When not present, its value shall be … Consider replacing "shall" with "is" similar to HEVC.]

**vuh\_reserved\_zero\_12bits**, when present, shall be equal to 0 in bitstreams conforming to this version of this document. Other values for vuh\_reserved\_zero\_12bits are reserved for future use by ISO/IEC. Decoders shall ignore the value of vuh\_reserved\_zero\_12bits.

**vuh\_reserved\_zero\_17bits**, when present, shall be equal to 0 in bitstreams conforming to this version of this document. Other values for vuh\_reserved\_zero\_17bits are reserved for future use by ISO/IEC. Decoders shall ignore the value of vuh\_reserved\_zero\_17bits.

**vuh\_reserved\_zero\_27bits**, when present, shall be equal to 0 in bitstreams conforming to this version of this document. Other values for vuh\_reserved\_zero\_27bits are reserved for future use by ISO/IEC. Decoders shall ignore the value of vuh\_reserved\_zero\_27bits.

#### V-PCC unit payload semantics

None.

#### Order of V-PCC units and association to coded information

A V-PCC bitstream contains a series of V-PCC sequences. A vpcc unit type with a value of vuh\_unit\_type equal to VPCC\_VPS is expected to be the first V-PCC unit type in a V-PCC sequence. All other V-PCC unit types shall follow this unit type without any additional restrictions in their coding order.

[Ed. (ALL): This section is very thin and needs to be extended to similar extent as the section for NAL unit order. This should also specify some constraints on the values of certain parameters and how they are allowed (or not) on how to change]

[Ed. (AMT): Unfortunately, the MIV spec does not define this correctly since it seems to confuse vpcc unit types with VPCC IRAP units. They are not the same thing. A VPCC unit type can contain an entire stream of a certain type. It is indeed a requirement that there should be at least certain VPCC unit types in the stream that relate to one another and that the defined set of those so as to allow the reconstruction of a point cloud, e.g. an atlas, occupancy, geometry, and attributes if available, shall all contain a corresponding IRAP type at the start of such unit. But that needs to be better described. ]

### Byte alignment semantics

**alignment\_bit\_equal\_to\_one** shall be equal to 1.

**alignment\_bit\_equal\_to\_zero** shall be equal to 0.

### V-PCC parameter set semantics

#### General V-PCC parameter set semantics

**vps\_vpcc\_parameter\_set\_id** provides an identifier for the V-PCC VPS for reference by other syntax elements. The value of vps\_vpcc\_parameter\_set\_id shall be in the range of 0 to 15, inclusive.

**vps\_reserved\_zero\_8bits**, when present, shall be equal to 0 in bitstreams conforming to this version of this document. Other values for vps\_reserved\_zero\_8bits are reserved for future use ISO/IEC. Decoders shall ignore the value of vps\_reserved\_zero\_8bits.

**vps\_atlas\_count\_minus1** plus 1 indicates the total number of supported atlases in the current bitstream. The value of vps\_atlas\_count\_minus1 shall be in the range of 0 to 63, inclusive.

**vps\_frame\_width**[ j ] indicates the V-PCC frame width in terms of integer luma samples for the atlas with index j. This frame width is the nominal width that is associated with all V-PCC components for the atlas with index j.

**vps\_frame\_height**[ j ] indicates the V-PCC frame height in terms of integer luma samples for the atlas with index j. This frame height is the nominal height that is associated with all V-PCC components for the atlas with index j.

**vps\_map\_count\_minus1**[ j ] plus 1 indicates the number of maps used for encoding the geometry and attribute data for the atlas with index j. vps\_map\_count\_minus1[ j ] shall be in the range of 0 to 15, inclusive.

**vps\_multiple\_map\_streams\_present\_flag**[ j ] equal to 0 indicates that all geometry or attribute maps for the atlas with index j are placed in a single geometry or attribute video stream, respectively. vps\_multiple\_map\_streams\_present\_flag[ j ] equal to 1 indicates that all geometry or attribute maps for the atlas with index j are placed in separate video streams. When vps\_multiple\_map\_streams\_present\_flag[ j ] is not present, its value shall be inferred to be equal to 0.

**vps\_map\_absolute\_coding\_enabled\_flag**[ j ][ i ] equal to 1 indicates that the geometry map with index i for the atlas with index j is coded without any form of map prediction. vps\_map\_absolute\_coding\_enabled\_flag[ j ][ i ]equal to 0 indicates that the geometry map with index i for the atlas with index j is first predicted from another, earlier coded map, prior to coding. If vps\_map\_absolute\_coding\_enabled\_flag[ j ][ i ] is not present, its value shall be inferred to be equal to 1.

**vps\_map\_predictor\_index\_diff**[ j ][ i ] is used to compute the predictor of the geometry map with index i for the atlas with index j when vps\_map\_absolute\_coding\_enabled\_flag[ j ][ i ] is equal to 0. More specifically, the map predictor index for map i shall be computed as:

VPCCMapPredictorIndex( i ) = (i – 1) – vps\_map\_predictor\_index\_diff[ j ][ i ] (7‑1)

The value of vps\_map\_predictor\_index\_diff[ j ][ i ] shall be in the range from 0 to i - 1, inclusive. When vps\_map\_predictor\_index\_diff[ j ][ i ] is not present, its value shall be inferred to be equal to 0.

**vps\_auxiliary\_video\_present\_flag**[ j ]equal to 1 indicates that auxiliary information for the atlas with index j, i.e. RAW or EOM patch data, may be stored in a separate video stream, refered to as the auxiliary video stream. vps\_auxiliary\_video\_present\_flag[ j ] equal to 0 indicates that auxiliary information for the atlas with index j shall not be stored in a separate video stream. When vps\_auxiliary\_video\_present\_flag[ j ] is not present, it is inferred to be equal to 0.

**vps\_extension\_present\_flag** equal to 1 specifies that the syntax element vps\_extension\_length is present in vpcc\_parameter\_set syntax structure. vps\_extension\_present\_flag equal to 0 specifies that syntax element vps\_extension\_length is not present. vps\_extension\_present\_flag shall be equal to 0 in bitstreams conforming to this version of this document.

**vps\_extension\_length\_minus1** plus 1 specifies the number of vps\_extension\_data\_byte elements that follow this syntax element.

**vps\_extension\_data\_byte** may have any value.

#### Profile, tier, and level semantics

**ptl\_tier\_flag** specifies the tier context for the interpretation of ptl\_level\_idc as specified in Annex A.

**ptl\_profile\_codec\_group\_idc** indicates the codec group profile component to which the CPCS conforms as specified in Annex A. Bitstreams shall not contain values of ptl\_profile\_codec\_group\_idc other than those specified in Annex A. Other values of ptl\_profile\_codec\_group\_idc are reserved for future use by ISO/IEC.

**ptl\_profile\_pcc\_toolset\_idc** indicates the toolset combination profile component to which the CPCS conforms as specified in Annex A. Bitstreams shall not contain values of ptl\_profile\_pcc\_toolset\_idc other than those specified in Annex A. Other values of ptl\_profile\_pcc\_toolset\_idc are reserved for future use by ISO/IEC.

**ptl\_profile\_reconstruction\_idc** indicates the reconstruction profile component to which the CPCS is recommended to conform to as specified in Annex A. This indication is only informative and decoders may select to use a different reconstruction profile than the one indicated in the bitstream. Bitstreams shall not contain values of ptl\_profile\_reconstruction\_idc other than those specified in Annex A. Other values of ptl\_profile\_reconstruction\_idc are reserved for future use by ISO/IEC.

**ptl**\_**reserved\_zero\_32bits**, when present, shall be equal to 0 in bitstreams conforming to this version of this document. Other values for ptl\_reserved\_zero\_32bits are reserved for future use by ISO/IEC. Decoders shall ignore the value of ptl\_reserved\_zero\_32bits.

**ptl\_level\_idc** indicates a level to which the CPCS conforms as specified in Annex A Bitstreams shall not contain values of ptl\_level\_idc other than those specified in Annex A. Other values of ptl\_level\_idc are reserved for future use by ISO/IEC.

#### Occupancy information semantics

**oi\_occupancy\_codec\_id**[ j ]indicates the identifier of the codec used to compress the occupancy map information for the atlas with index j. occupancy\_codec\_id shall be in the range of 0 to 255, inclusive. This codec may be identified through a component codec mapping SEI message or through means outside this document.

**oi\_lossy\_occupancy\_map\_compression\_threshold**[ j ] indicates the threshold to be used to derive the binary occupancy map from the decoded occupancy map video for the atlas with index j. oi\_lossy\_occupancy\_map\_compression\_threshold[ j ] shall be in the range of 0 to 255, inclusive.

**oi\_occupancy\_nominal\_2d\_bitdepth\_minus1**[ j ] plus 1 indicates the nominal 2D bit depth to which the occupancy video for the atlas with index j shall be converted to. oi\_occupancy\_nominal\_2d\_bitdepth\_minus1[ j ] shall be in the range of 0 to 31, inclusive.

**oi\_occupancy\_MSB\_align\_flag**[ j ] indicates how the decoded occupancy video samples associated with an atlas with index j are converted to samples at the nominal occupancy bitdepth, as specified in Section B.2.1.

#### Geometry information semantics

**gi\_geometry\_codec\_id**[ j ]indicates the identifier of the codec used to compress the geometry video data for the atlas with index j. geometry\_codec\_id shall be in the range of 0 to 255, inclusive. This codec may be identified through a component codec mapping SEI message or through means outside this document.

**gi\_geometry\_nominal\_2d\_bitdepth\_minus1**[ j ] plus 1 indicates the nominal 2D bit depth to which all geometry videos for the atlas with index j shall be converted to. gi\_geometry\_nominal\_2d\_bitdepth\_minus1[ j ] shall be in the range of 0 to 31, inclusive.

**gi\_geometry\_MSB\_align\_flag**[ j ] indicates how the decoded geometry video samples associated with an atlas with index j are converted to samples at the nominal geometry bitdepth, as specified in Section B.2.1.

**gi\_geometry\_3d\_coordinates\_bitdepth\_minus1**[ j ] plus 1 indicates the bit depth of the geometry coordinates of the reconstructed point cloud for the atlas with index j. gi\_geometry\_3d\_coordinates\_bitdepth\_minus1[ j ]shall be in the range of 0 to 31, inclusive.

**gi\_auxiliary\_geometry\_codec\_id**[ j ], when present,indicates the identifier of the codec used to compress the geometry video data sub-bitstreams, when RAW coded points are encoded in an auxiliary video stream for the atlas with index j. gi\_auxiliary\_geometry\_codec\_id[ j ] shall be in the range of 0 to 255, inclusive. This codec may be identified through a component codec mapping SEI message or through means outside this document. When not present the value of gi\_auxiliary\_geometry\_codec\_id[ j ] shall be set equal to gi\_geometry\_codec\_id[ j ].

#### Attribute information semantics

**ai\_attribute\_count**[ j ]indicates the number of attributes associated with the point cloud for the atlas with index j. vpcc\_attribute\_count shall be in the range of 0 to 127, inclusive.

**ai\_attribute\_type\_id**[ j ][ i ] indicates the attribute type of the Attribute Video Data unit with index i for the atlas with index j. Table 7‑2 describes the list of supported attributes and their relationship with ai\_attribute\_type\_id.

Table 7‑2 V-PCC attribute types

|  |  |  |
| --- | --- | --- |
| **ai\_attribute\_type\_id**[ j ][ i ] | **Identifier** | **Attribute type** |
| 0 | ATTR\_TEXTURE | Texture |
| 1 | ATTR\_MATERIAL\_ID | Material ID |
| 2 | ATTR\_TRANSPARENCY | Transparency |
| 3 | ATTR\_REFLECTANCE | Reflectance |
| 4 | ATTR\_NORMAL | Normals |
| 5…14 | ATTR\_RESERVED | Reserved |
| 15 | ATTR\_UNSPECIFIED | Unspecified |

ATTR\_TEXTURE indicates an attribute that contains texture information of a point cloud. For example, this may indicate an attribute that contains RGB (Red, Green, Blue) colour information.

ATTR\_MATERIAL\_ID indicates an attribute that contains supplemental information that indicates the material type of a point in a point cloud. For example, the material type could be used as a indicator for identifying an object or the characteristic of a point within a point cloud. The interpretation of the values of such attribute frame type is outside the scope of this document.

ATTR\_TRANSPARENCY indicates an attribute that contains transparency information that is associated with each point in a point cloud.

ATTR\_REFLECTANCE indicates an attribute that contains reflectance information that is associated with each point in a point cloud.

ATTR\_NORMAL indicates an attribute that contains a unit vector information associated with each point in a point cloud. The unit vector specifies the perpendicular direction to a surface at a point (i.e. direction a point is facing). An attribute frame with this attribute type shall have ai\_attribute\_dimension\_minus1 equal to 2. Each channel of an attribute frame with this attribute type shall contain one component of the unit vector (x, y, z), where the first component contains the x coordinate, the second component contains the y coordinate, and the third component contains the z coordinate.

ATTR\_UNSPECIFIED indicates an attribute that contains values that have no specified meaning in this document and will not have a specified meaning in the future as an integral part of this document.

**ai\_attribute\_codec\_id**[ j ][ i ]indicates the identifier of the codec used to compress the attribute video data with index i for the atlas with index j. ai\_attribute\_codec\_id[ j ][ i ] shall be in the range of 0 to 255, inclusive. This codec may be identified through a component codec mapping SEI message or through means outside this document.

**ai\_auxiliary\_attribute\_codec\_id**[ j ][ i ], when present,indicates the identifier of the codec used to compress the attribute video data for RAW and/or EOM coded points of attribute i, when RAW and/or EOM coded points are encoded in an auxiliary video stream for the atlas with index j. ai\_auxiliary\_attribute\_codec\_id[ j ][ i ] shall be in the range of 0 to 255, inclusive. This codec may be identified through a component codec mapping SEI message or through means outside this document. When not present the value of ai\_auxiliary\_attribute\_codec\_id[ j ][ i ] shall be set equal to ai\_attribute\_codec\_id[ j ][ i ].

**ai\_attribute\_map\_absolute\_coding\_persistence\_flag**[ k ][ j ] equal to 1 indicates that all attribute maps, for the attribute with index j, that corresponds to the atlas with index k, are coded without any form of map prediction. ai\_attribute\_map\_absolute\_coding\_persistence\_flag[ k ][ j ] equal to 0 indicates that the attribute maps of the attribute with index j, that correspond to the atlas with index k, will use the same map prediction method as used for the geometry of atlas with index k. If ai\_attribute\_map\_absolute\_coding\_persistence\_flag[ k ][ j ] is not present, its value shall be inferred to be equal to 1.

Then the 3D array AttributeMapAbsoluteCodingEnabledFlag, which indicates if a particular map of an attribute is to be coded with or without prediction, is obtained as follows:

if( ai\_attribute\_map\_absolute\_coding\_persistance\_flag[ atlasId ][ i ] == 1) {

for( j = 0; j < vps\_map\_count\_minus1[ atlasId ]; j++ )

AttributeMapAbsoluteCodingEnabledFlag[ atlasId ][ i ][ j ] = 1

}

else{

for( j = 0; j < vps\_map\_count\_minus1[ atlasId ]; j++ )

AttributeMapAbsoluteCodingEnabledFlag[ atlasId ][ i ][ j ] =

vps\_map\_absolute\_coding\_enabled\_flag[ atlasId ][ j ]

}

**ai\_attribute\_dimension\_minus1**[ j ][ i ] plus 1 indicates the total number of dimensions (i.e., number of channels) of the attribute with index i for the atlas with index j. ai\_attribute\_dimension\_minus1[ j ][ i ]shall be in the range of 0 to 63, inclusive.

**ai\_attribute\_dimension\_partitions\_minus1**[ j ][ i ] plus 1 indicates the number of partition groups in which the attribute channels for attribute with index i for the atlas with index j should be grouped in. ai\_attribute\_dimension\_partitions\_minus1[ j ][ i ]shall be in the range of 0 to 63, inclusive.

**ai\_attribute\_partition\_channels\_minus1**[ k ][ i ][ j ] plus 1 indicates the number of channels assigned to the dimension partition group with index j of attribute with index i for atlas with index k. ai\_attribute\_partition\_channels\_minus1[ k ][ i ][ j ]shall be in the range of 0 to ai\_attribute\_dimension\_minus1[ k ][ i ], inclusive, for all dimension partition groups.

**ai\_attribute\_nominal\_2d\_bitdepth\_minus1**[ j ][ i ] plus 1 indicates the nominal 2D bit depth to which all the attribute videos with attribute index i for atlas with index j shall be converted to. ai\_attribute\_nominal\_2d\_bitdepth\_minus1[ j ][ i ] shall be in the range of 0 to 31, inclusive.

**ai\_attribute\_MSB\_align\_flag**[ j ][ i ] indicates how the decoded attribute video samples, associated with an atlas with index j and attribute with index i, are converted to samples at the nominal attribute bitdepth, as specified in Section B.2.1.

[Ed. (RJ): It is not clear whether the conversion should be applied after the decoded attribute frame has been converted to the nominal resolution or immediately after decoding the attribute video. This needs to be specified for conformance point B.]

[Ed. (AMT): Why is this set the same for all attributes? It might be the case that for colours we use one method but for material id we use a different one. This seems too restrictive. The cost also of sending these seems small since this are binary values. This may, for example, impact coding of material ID information.]

### NAL unit semantics

#### General NAL unit semantics

NumBytesInNalUnit specifies the size of the NAL unit in bytes. This value is required for decoding of the NAL unit. Some form of demarcation of NAL unit boundaries is necessary to enable inference of NumBytesInNalUnit. One such demarcation method is specified in Annex C for the sample stream format. Other methods of demarcation may be specified outside of this document.

NOTE 1 – The atlas coding layer (ACL) is specified to efficiently represent the content of the patch data. The NAL is specified to format that data and provide header information in a manner appropriate for conveyance on a variety of communication channels or storage media. All data are contained in NAL units, each of which contains an integer number of bytes. A NAL unit specifies a generic format for use in both packet-oriented and bitstream systems. The format of NAL units for both packet-oriented transport and sample streams is identical except that in the sample stream format specified in Annex C each NAL unit can be preceded by an additional element that specifies the size of the NAL unit.

**rbsp\_byte**[ i ] is the i-th byte of an RBSP. An RBSP is specified as an ordered sequence of bytes as follows:

The RBSP contains a string of data bits(SODB) as follows:

– If the SODB is empty (i.e., zero bits in length), the RBSP is also empty.

– Otherwise, the RBSP contains the SODB as follows:

1) The first byte of the RBSP contains the first (most significant, left-most) eight bits of the SODB; the next byte of the RBSP contains the next eight bits of the SODB, etc., until fewer than eight bits of the SODB remain.

2) The rbsp\_trailing\_bits( ) syntax structure is present after the SODB as follows:

i) The first (most significant, left-most) bits of the final RBSP byte contain the remaining bits of the SODB (if any).

ii) The next bit consists of a single bit equal to 1 (i.e., rbsp\_stop\_one\_bit).

iii) When the rbsp\_stop\_one\_bit is not the last bit of a byte-aligned byte, one or more bits equal to 0 (i.e. instances of rbsp\_alignment\_zero\_bit) are present to result in byte alignment.

3) One or more cabac\_zero\_word 16-bit syntax elements equal to 0x0000 may be present in some RBSPs after the rbsp\_trailing\_bits( ) at the end of the RBSP.

Syntax structures having these RBSP properties are denoted in the syntax tables using an "\_rbsp" suffix. These structures are carried within NAL units as the content of the rbsp\_byte[ i ] data bytes. The association of the RBSP syntax structures to the NAL units is as specified in Table 7‑3.

NOTE 2 – When the boundaries of the RBSP are known, the decoder can extract the SODB from the RBSP by concatenating the bits of the bytes of the RBSP and discarding the rbsp\_stop\_one\_bit, which is the last (least significant, right-most) bit equal to 1, and discarding any following (less significant, farther to the right) bits that follow it, which are equal to 0. The data necessary for the decoding process is contained in the SODB part of the RBSP.

#### NAL unit header semantics

**nal\_forbidden\_zero\_bit** shall be equal to 0.

**nal\_unit\_type** specifies the type of the RBSP data structure contained in the NAL unit as specified in Table 7‑3.

NAL units that have nal\_unit\_type in the range of NAL\_UNSPEC\_48..NAL\_UNSPEC\_63, inclusive, for which semantics are not specified, shall not affect the decoding process specified in this document.

NOTE 1 – NAL unit types in the range of NAL\_UNSPEC\_48..NAL\_UNSPEC\_63 may be used as determined by the application. No decoding process for these values of nal\_unit\_type is specified in this document. Since different applications might use these NAL unit types for different purposes, particular care must be exercised in the design of encoders that generate NAL units with these nal\_unit\_type values, and in the design of decoders that interpret the content of NAL units with these nal\_unit\_type values. This document does not define any management for these values. These nal\_unit\_type values might only be suitable for use in contexts in which "collisions" of usage (i.e., different definitions of the meaning of the NAL unit content for the same nal\_unit\_type value) are unimportant, or not possible, or are managed – e.g., defined or managed in the controlling application or transport specification, or by controlling the environment in which bitstreams are distributed.

For purposes other than determining the amount of data in the decoding units of the bitstream (as specified in Annex C), decoders shall ignore (i.e. remove from the bitstream and discard) the contents of all NAL units that use reserved values of nal\_unit\_type.

NOTE 2 – This requirement allows future definition of compatible extensions to this document.

Table 7‑3 – NAL unit type codes and NAL unit type classes

|  |  |  |  |
| --- | --- | --- | --- |
| **nal\_unit\_type** | **Name of nal\_unit\_type** | **Content of NAL unit and RBSP syntax structure** | **NAL unit type class** |
| 0 | NAL\_TRAIL | Coded tile group of a non-TSA, non STSA trailing atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 1 | NAL\_TSA | Coded tile group of a TSA atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 2 | NAL\_STSA | Coded tile group of a STSA atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 3 | NAL\_RADL | Coded tile group of a RADL atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 4 | NAL\_RASL | Coded tile group of a RASL atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 5 | NAL\_SKIP | Coded tile group of a skipped atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 6..9 | NAL\_RSV\_ACL\_6.. NAL\_RSV\_ACL\_9 | Reserved non-IRAP ACL NAL unit types | ACL |
| 10 11 12 | NAL\_BLA\_W\_LP NAL\_BLA\_W\_RADL NAL\_BLA\_N\_LP | Coded tile group of a BLA atlas frame atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 13 14 15 | NAL\_GBLA\_W\_LP NAL\_GBLA\_W\_RADL NAL\_GBLA\_N\_LP | Coded tile group of a GBLA atlas frame atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 16 17 | NAL\_IDR\_W\_RADL NAL\_IDR\_N\_LP | Coded tile group of an IDR atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 18 19 | NAL\_GIDR\_W\_RADL NAL\_GIDR\_N\_LP | Coded tile group of a GIDR atlas frame  atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 20 | NAL\_CRA | Coded tile group of a CRA atlas frame atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 21 | NAL\_GCRA | Coded tile group of a GCRA atlas frame atlas\_tile\_group\_layer\_rbsp( ) | ACL |
| 22 23 | NAL\_IRAP\_ACL\_22 NAL\_IRAP\_ACL\_23 | Reserved IRAP ACL NAL unit types | ACL |
| 24..31 | NAL\_RSV\_ACL\_24.. NAL\_RSV\_ACL\_31 | Reserved non-IRAP ACL NAL unit types | ACL |
| 32 | NAL\_ASPS | Atlas sequence parameter set atlas\_sequence\_parameter\_set\_rbsp( ) | non-ACL |
| 33 | NAL\_AFPS | Atlas frame parameter set atlas\_frame\_parameter\_set\_rbsp( ) | non-ACL |
| 34 | NAL\_AUD | Access unit delimiter access\_unit\_delimiter\_rbsp( ) | non-ACL |
| 35 | NAL\_VPCC\_AUD | V-PCC access unit delimiter access\_unit\_delimiter\_rbsp( ) | non-ACL |
| 36 | NAL\_EOS | End of sequence end\_of\_seq\_rbsp( ) | non-ACL |
| 37 | NAL\_EOB | End of bitstream end\_of\_atlas\_sub\_bitstream\_rbsp( ) | non-ACL |
| 38 | NAL\_FD | Filler filler\_data\_rbsp( ) | non-ACL |
| 39 40 | NAL\_PREFIX\_NSEI  NAL\_SUFFIX\_NSEI | Non-essential supplemental enhancement information sei\_rbsp( ) | non-ACL |
| 41 42 | NAL\_PREFIX\_ESEI NAL\_SUFFIX\_ESEI | Essential supplemental enhancement information sei\_rbsp( ) | non-ACL |
| 43 | NAL\_AAPS | Atlas adaptation parameter set atlas\_adaptation\_parameter\_set\_rbsp( ) | non-ACL |
| 44..47 | NAL\_RSV\_NACL\_44 NAL\_RSV\_NACL\_47 | Reserved non-ACL NAL unit types | non-ACL |
| 48..63 | NAL\_UNSPEC\_48 NAL\_UNSPEC\_63 | Unspecified non-ACL NAL unit types | non-ACL |

NOTE 3 – A clean random access (CRA) and a global clean random access atlas frame may have associated random access skipped leading (RASL) or random access decodable leading (RADL) atlas frames present in the bitstream.

NOTE 4 – A broken link access (BLA) or a global broken link access (GBLA) atlas frame having nal\_unit\_type equal to NAL\_BLA\_W\_LP and NAL\_GBLA\_W\_LP respectively, may have associated RASL or RADL atlas frames present in the bitstream. A BLA or a GBLA atlas frame having nal\_unit\_type equal to NAL\_BLA\_W\_RADL and NAL\_GBLA\_W\_RADL respectively, does not have associated RASL atlas frames present in the bitstream, but may have associated RADL atlas frames in the bitstream. A BLA or a GBLA atlas frame having nal\_unit\_type equal to NAL\_BLA\_N\_LP and NAL\_GBLA\_N\_LP respectively, does not have associated leading atlas frames present in the bitstream.

NOTE 5 – An instantaneous decoding refresh (IDR) or a global instantaneous decoding refresh (GIDR) atlas frame having nal\_unit\_type equal to NAL\_IDR\_N\_LP and NAL\_GIDR\_N\_LP respectively, does not have associated leading atlas frames present in the bitstream. An IDR or a GIDR atlas frame having nal\_unit\_type equal to NAL\_IDR\_W\_RADL and NAL\_GIDR\_W\_RADL respectively, does not have associated RASL atlas frames present in the bitstream, but may have associated RADL atlas frames in the bitstream.

NOTE 6 – A sub-layer non-reference (SLNR) atlas frame is not included in any of RefPicSetStCurrBefore, RefPicSetStCurrAfter and RefPicSetLtCurr of any atlas frame with the same value of TemporalId, and may be discarded without affecting the decodability of other atlas frames with the same value of TemporalId.

All coded tile group NAL units of an access unit shall have the same value of nal\_unit\_type. An atlas frame or an access unit is also referred to as having a nal\_unit\_type equal to the nal\_unit\_type of the coded tile group NAL units of the atlas frame or the access unit.

[Ed. (LK): The above NAL units although may seem to provide certain functionality were not explicitly discussed or studied during the last meeting. It is recommended that these should be discussed further during the upcoming meeting before adopting them in the specification]

If an atlas frame has nal\_unit\_type equal to NAL\_TRAIL, NAL\_TSA, NAL\_STSA, NAL\_RADL, and NAL\_RASL, the atlas frame is an SLNR atlas frame. Otherwise, the atlas frame is a sub-layer reference atlas frame.

Each atlas frame, other than the first atlas frame in the bitstream in decoding order, is considered to be associated with the previous intra random access point (IRAP) atlas frame in decoding order.

When an atlas frame is a leading atlas frame, it shall be a RADL or RASL atlas frame.

When an atlas frame is a trailing atlas frame, it shall not be a RADL or RASL atlas frame.

When an atlas frame is a leading atlas frame, it shall precede, in decoding order, all trailing atlas frames that are associated with the same IRAP atlas frame.

No RASL atlas frames shall be present in the bitstream that are associated with a BLA atlas frame having nal\_unit\_type equal to NAL\_BLA\_W\_RADL or NAL\_BLA\_N\_LP.

No RASL atlas frames shall be present in the bitstream that are associated with a GBLA atlas frame having nal\_unit\_type equal to NAL\_GBLA\_W\_RADL or NAL\_GBLA\_N\_LP.

No RASL atlas frames shall be present in the bitstream that are associated with an IDR or a GIDR atlas frame.

No RADL atlas frames shall be present in the bitstream that are associated with a BLA atlas frame having nal\_unit\_type equal to NAL\_BLA\_N\_LP, a GBLA atlas frame having nal\_unit\_type equal to NAL\_GBLA\_N\_LP, an IDR atlas frame having nal\_unit\_type equal to NAL\_IDR\_N\_LP, or a GIDR atlas frame having nal\_unit\_type equal to NAL\_GIDR\_N\_LP.

NOTE 7 – It is possible to perform random access at the position of an IRAP access unit by discarding all access units before the IRAP access unit (and to correctly decode the IRAP atlas frame and all the subsequent non-RASL atlas frames in decoding order), provided each parameter set is available (either in the bitstream or by external means not specified in this document) when it needs to be activated.

Any RASL atlas frame associated with a CRA, GCRA, BLA, or GBLA atlas frame shall precede any RADL atlas frame associated with the CRA, GCRA, BLA, or GBLA atlas frame in output order.

Any RASL atlas frame associated with a CRA or GCRA atlas frame shall follow, in output order, any IRAP atlas frame that precedes the CRA or GCRA atlas frame in decoding order.

A nal\_unit\_type equal to NAL\_VPCC\_AUD that is associated with a particular atlas frame j means that all output V-PCC component frames that have the same output time order as the atlas frame j shall have the same decoding order as that of the atlas frame.

NOTE 8 – An application may specify that all frames in a V-PCC sequence are delimited using a nal\_unit\_type equal to NAL\_VPCC\_AUD, in which case all V-PCC sub-bitstreams including sub-bitstreams that correspond to different maps will be aligned in both decoding order and output time.

**nal\_layer\_id** specifies the identifier of the layer to which an ACL NAL unit belongs or the identifier of a layer to which a non-ACL NAL unit applies. The value of nal\_layer\_id shall be in the range of 0 to 62, inclusive. The value of 63 may be specified in the future by ISO/IEC. For purposes other than determining the amount of data in the decoding units of the bitstream, decoders shall ignore all data that follow the value 63 for nal\_layer\_id in a NAL unit, and decoders conforming to a profile specified in Annex A shall ignore (i.e., remove from the bitstream and discard) all NAL units with values of nal\_layer\_id not equal to 0.

NOTE 9 – The value of 63 for nal\_layer\_id may be used to indicate an extended layer identifier in a future extension of this document.

The value of nal\_layer\_id shall be the same for all ACL NAL units of a coded atlas frame. The value of nal\_layer\_id of a coded atlas frame is the value of the nal\_layer\_id of the ACL NAL units of the coded atlas frame.

When nal\_unit\_type is equal to NAL\_EOB, the value of nal\_layer\_id shall be equal to 0.

**nal\_temporal\_id\_plus1** minus 1 specifies a temporal identifier for the NAL unit. The value of nal\_temporal\_id\_plus1 shall not be equal to 0.

The variable TemporalId is specified as follows:

TemporalId = nal\_temporal\_id\_plus1 − 1 (7‑2)

When nal\_unit\_type is in the range of NAL\_BLA\_W\_LP to NAL\_RSV\_IRAP\_ACL23, inclusive, i.e., the coded tile group belongs to an IRAP atlas frame, TemporalId shall be equal to 0.

When nal\_unit\_type is equal to NAL\_TSA, TemporalId shall not be equal to 0.

When nal\_layer\_id is equal to 0 and nal\_unit\_type is equal to NAL\_STSA, TemporalId shall not be equal to 0.

The value of TemporalId shall be the same for all ACL NAL units of an access unit. The value of TemporalId of a coded atlas frame or an access unit is the value of the TemporalId of the ACL NAL units of the coded atlas frame or the access unit. The value of TemporalId of a sub-layer representation is the greatest value of TemporalId of all ACL NAL units in the sub-layer representation.

The value of TemporalId for non-ACL NAL units is constrained as follows:

– If nal\_unit\_type is equal to NAL\_ASPS, TemporalId shall be equal to 0 and the TemporalId of the access unit containing the NAL unit shall be equal to 0.

– Otherwise if nal\_unit\_type is equal to NAL\_EOS or NAL\_EOB, TemporalId shall be equal to 0.

– Otherwise, if nal\_unit\_type is equal to NAL\_AUD, NAL\_VPCC\_AUD, or NAL\_FD, TemporalId shall be equal to the TemporalId of the access unit containing the NAL unit.

– Otherwise, TemporalId shall be greater than or equal to the TemporalId of the access unit containing the NAL unit.

NOTE 10 – When the NAL unit is a non-ACL NAL unit, the value of TemporalId is equal to the minimum value of the TemporalId values of all access units to which the non-ACL NAL unit applies. When nal\_unit\_type is equal to NAL\_AFPS, TemporalId may be greater than or equal to the TemporalId of the containing access unit, as all atlas frame parameter sets (AFPSs) may be included in the beginning of a bitstream, wherein the first coded atlas frame has TemporalId equal to 0. When nal\_unit\_type is equal to NAL\_PREFIX\_NSEI, NAL\_PREFIX\_ESEI, NAL\_SUFFIX\_NSEI, or NAL\_SUFFIX\_ESEI, TemporalId may be greater than or equal to the TemporalId of the containing access unit, as an SEI NAL unit may contain information, e.g., in a buffering period SEI message or an atlas frame timing SEI message, that applies to a bitstream subset that includes access units for which the TemporalId values are greater than the TemporalId of the access unit containing the SEI NAL unit.

[Ed. (LK): The above elements were included in the related contributions for adding NAL units in V-PCC however were not thoroughly discussed. Should these be discussed more carefully at the upcoming meeting, including their functionality/characteristics? Furthermore, if these are to be retained we should make sure that profiles and levels set these properly.]

[Ed. AMT: Do we need to add any TemporalID considerations for the adaptation parameter sets]

#### Order of NAL units and association to coded atlas frames, access units, and coded atlas sequences

##### General

This clause specifies constraints on the order of NAL units in the bitstream.

Any order of NAL units in the bitstream obeying these constraints is referred to in the text as the decoding order of NAL units. Within a NAL unit, the syntax in clauses 7.3, specifies the decoding order of syntax elements. Decoders shall be capable of receiving NAL units and their syntax elements in decoding order.

##### Order of AAPS, ASPS, and AFPS RBSPs and their activation

This clause specifies the activation process of atlas adaptation parameter sets (AAPSs), atlas sequence parameter sets (ASPSs), and atlas frame parameter sets (AFPSs).

NOTE – The AAPS, ASPS, and AFPS mechanism decouples the transmission of infrequently changing information from the transmission of coded atlas data. AAPSs, ASPSs, and AFPSs may, in some applications, be conveyed "out-of-band".

An AAPS RBSP includes parameters that can be referred to by the coded tile group NAL units of one or more coded atlas frames. Each AAPS RBSP is initially considered not active at the start of the operation of the decoding process. At most one AAPS RBSP is considered active at any given moment during the operation of the decoding process, and the activation of any particular AAPS RBSP results in the deactivation of the previously-active AAPS RBSP.

When an AAPS RBSP (with a particular value of aaps\_atlas\_adaptation\_parameter\_set\_id) is not active and is referred to by a coded tile group NAL unit with nal\_layer\_id equal to 0 (using a value of atgh\_atlas\_adapatation\_parameter\_set\_id equal to the aaps\_atlas\_activation\_parameter\_set\_id value), it is then activated. This AAPS RBSP is called the active AAPS RBSP until it is deactivated by the activation of another AAPS RBSP. An AAPS RBSP, with that particular value of aaps\_atlas\_activation\_parameter\_set\_id, shall be available to the decoding process prior to its activation, included in at least one access unit with TemporalId less than or equal to the TemporalId of the AAPS NAL unit or provided through external means, and the AAPS NAL unit containing the AAPS RBSP shall have nal\_layer\_id equal to 0.

[Ed. (AMT): Do we wish to permit an AAPS to be missing from the stream? One could argue that this is already handled by allowing AAPS to be conveyed out of band]

Any AAPS NAL unit containing the value of aaps\_atlas\_adaptation\_parameter\_set\_id for the active AAPS RBSP for a coded atlas frame shall have the same content as that of the active AAPS RBSP for the coded atlas frame, unless it follows the last ACL NAL unit of the coded atlas frame and precedes the first ACL NAL unit of another coded atlas frame.

An AFPS RBSP includes parameters that can be referred to by the coded tile group NAL units of one or more coded atlas frames. Each AFPS RBSP is initially considered not active at the start of the operation of the decoding process. At most one AFPS RBSP is considered active at any given moment during the operation of the decoding process, and the activation of any particular AFPS RBSP results in the deactivation of the previously-active AFPS RBSP.

When an AFPS RBSP (with a particular value of afps\_atlas\_frame\_parameter\_set\_id) is not active and is referred to by a coded tile group NAL unit with nal\_layer\_id equal to 0 (using a value of atgh\_atlas\_frame\_parameter\_set\_id equal to the afps\_atlas\_frame\_parameter\_set\_id value), it is then activated. This AFPS RBSP is called the active AFPS RBSP until it is deactivated by the activation of another AFPS RBSP. An AFPS RBSP, with that particular value of afps\_atlas\_frame\_parameter\_set\_id, shall be available to the decoding process prior to its activation, included in at least one access unit with TemporalId less than or equal to the TemporalId of the AFPS NAL unit or provided through external means, and the AFPS NAL unit containing the AFPS RBSP shall have nal\_layer\_id equal to 0.

Any AFPS NAL unit containing the value of afps\_atlas\_frame\_parameter\_set\_id for the active AFPS RBSP for a coded atlas frame shall have the same content as that of the active AFPS RBSP for the coded atlas frame, unless it follows the last ACL NAL unit of the coded atlas frame and precedes the first ACL NAL unit of another coded atlas frame.

An ASPS RBSP includes parameters that can be referred to by one or more AFPS RBSPs. Each ASPS RBSP is initially considered not active at the start of the operation of the decoding process. At most one ASPS RBSP is considered active at any given moment during the operation of the decoding process, and the activation of any particular ASPS RBSP results in the deactivation of the previously-active ASPS RBSP.

When an ASPS RBSP (with a particular value of asps\_atlas\_sequence\_parameter\_set\_id) is not already active and it is referred to by activation of a AFPS RBSP (in which afps\_seq\_parameter\_set\_id is equal to the asps\_atlas\_sequence\_parameter\_set\_id value), it is activated. This ASPS RBSP is called the active ASPS RBSP until it is deactivated by the activation of another ASPS RBSP. An ASPS RBSP, with that particular value of asps\_atlas\_sequence\_parameter\_set\_id, shall be available to the decoding process prior to its activation, included in at least one access unit with TemporalId equal to 0 or provided through external means, and the NAL unit containing the ASPS RBSP shall have nal\_layer\_id equal to 0. An activated ASPS RBSP shall remain active for the entire coded atlas sequence (CAS).

Any ASPS NAL unit with nal\_layer\_id equal to 0 containing the value of asps\_atlas\_sequence\_parameter\_set\_id for the active ASPS RBSP for a CAS shall have the same content as that of the active ASPS RBSP for the CAS, unless it follows the last access unit of the CAS and precedes the first ACL NAL unit of another CAS.

All constraints that are expressed on the relationship between the values of the syntax elements and the values of variables derived from those syntax elements in AAPSs, ASPSs, and AFPSs and other syntax elements are expressions of constraints that apply only to the active AAPS RBSP, the active ASPS RBSP, and the active AFPS RBSP. If any AAPS RBSP, ASPS RBSP, and AFPS RBSP is present that is never activated in the bitstream, its syntax elements shall have values that would conform to the specified constraints if it was activated by reference in an otherwise conforming bitstream.

During operation of the decoding process (see clause 8), the values of parameters of the active AAPS RBSP, the active ASPS RBSP, and the active AFPS RBSP are considered in effect. For interpretation of SEI messages, the values of the active AAPS RBSP, the active ASPS RBSP, and the active AFPS RBSP for the operation of the decoding process for the ACL NAL units of the coded atlas frame with nal\_layer\_id equal to 0 in the same access unit are considered in effect unless otherwise specified in the SEI message semantics.

##### Order of access units and association to CASs

An bitstream conforming to this document consists of one or more CASs.

A CAS consists of one or more access units. The order of NAL units and coded atlas frames and their association to access units is described in clause 7.4.5.3.4.

The first access unit of a CAS is an IRAP access unit with NoRaslOutputFlag equal to 1.

It is a requirement of bitstream conformance that, when present, the next access unit after an access unit that contains an end of sequence NAL unit or an end of an bitstream NAL unit shall be an IRAP access unit, which may be an IDR access unit, a GIDR access unit, a BLA access unit, a GBLA access unit, a CRA access unit, or a GCRA access unit.

##### Order of NAL units and coded atlas frames and their association to access units

This clause specifies the order of NAL units and coded atlas frames and their association to access units for CPCSs that conform to one or more of the profiles specified in Annex A and that are decoded using the decoding process specified in clauses 2 through 9.6.

An access unit consists of one coded atlas with nal\_layer\_id equal to 0, zero or more ACL NAL units with nal\_layer\_id greater than 0 and zero or more non-ACL NAL units. The association of ACL NAL units to coded atlases is described in clause 7.4.5.3.5.

The first access unit in the bitstream starts with the first NAL unit of the bitstream.

Let firstBlAFrmNalUnit be the first ACL NAL unit of a coded atlas frame with nal\_layer\_id equal to 0. The first of any of the following NAL units preceding firstBlAFrmNalUnit and succeeding the last ACL NAL unit preceding firstBlAFrmNalUnit, if any, specifies the start of a new access unit:

NOTE 1 – The last ACL NAL unit preceding firstBlAFrmNalUnit in decoding order may have nal\_layer\_id greater than 0.

– access or V-PCC access unit delimiter NAL unit with nal\_layer\_id equal to 0 (when present),

– ASPS NAL unit with nal\_layer\_id equal to 0 (when present),

– AFPS NAL unit with nal\_layer\_id equal to 0 (when present),

– Prefix SEI NAL unit with nal\_layer\_id equal to 0 (when present),

– NAL units with nal\_unit\_type in the range of NAL\_RSV\_NACL41..NAL\_RSV\_NACL44 with nal\_layer\_id equal to 0 (when present),

– NAL units with nal\_unit\_type in the range of NAL\_UNSPEC\_48..NAL\_UNSPEC\_55 with nal\_layer\_id equal to 0 (when present).

NOTE 2 – The first NAL unit preceding firstBlAFrmNalUnit and succeeding the last ACL NAL unit preceding firstBlAFrmNalUnit, if any, can only be one of the above-listed NAL units.

When there is none of the above NAL units preceding firstBlAFrmNalUnit and succeeding the last ACL NAL preceding firstBlAFrmNalUnit, if any, firstBlAFrmNalUnit starts a new access unit.

The order of the coded atlas frames and non-ACL NAL units within an access unit shall obey the following constraints:

– When an access unit delimiter or a V-PCC access unit delimiter NAL unit with nal\_layer\_id equal to 0 is present, it shall be the first NAL unit. There shall be at most one access unit delimiter or V-PCC access unit delimiter NAL unit with nal\_layer\_id equal to 0 in any access unit.

– When any ASPS NAL units, AFPS NAL units, prefix SEI NAL units, NAL units with nal\_unit\_type in the range of NAL\_RSV\_NACL\_41..NAL\_RSV\_NACL\_44, or NAL units with nal\_unit\_type in the range of NAL\_UNSPEC\_48..NAL\_UNSPEC\_55 are present, they shall not follow the last ACL NAL unit of the access unit.

– NAL units having nal\_unit\_type equal to NAL\_FDT, NAL\_SUFFIX\_NSEI, or NAL\_SUFFIX\_ESEI or in the range of NAL\_RSV\_NACL\_45..NAL\_RSV\_NACL\_47 or NAL\_UNSPEC\_56..NAL\_UNSPEC\_63 shall not precede the first ACL NAL unit of the access unit.

– When an end of sequence NAL unit with nal\_layer\_id equal to 0 is present, it shall be the last NAL unit among all NAL units with nal\_layer\_id equal to 0 in the access unit other than an end of an bitstream NAL unit (when present).

– When an end of an bitstream NAL unit is present, it shall be the last NAL unit in the access unit.

NOTE 3 – Decoders conforming to profiles specified in Annex A do not use NAL units with nal\_layer\_id greater than 0, e.g., access or V-PCC access unit delimiter NAL units with nal\_layer\_id greater than 0, for access unit boundary detection, except for identification of a NAL unit as an ACL or non-ACL NAL unit.

##### Order of ACL NAL units and association to coded atlas frames

This clause specifies the order of ACL NAL units and association to coded atlas frames.

Each ACL NAL unit is part of a coded atlas frame.

The order of the ACL NAL units within a coded atlas frame is constrained as follows:

– The first ACL NAL unit of the coded atlas frame shall have atgh\_address equal to 0.

### Raw byte sequence payloads, trailing bits, and byte alignment semantics

#### Atlas sequence parameter set RBSP semantics

##### General atlas sequence parameter set RBSP semantics

**asps\_atlas\_sequence\_parameter\_set\_id** provides an identifier for the atlas sequence parameter set for reference by other syntax elements. The value of asps\_atlas\_sequence\_parameter\_set\_id shall be in the range of 0 to 15, inclusive.

**asps\_frame\_width** indicates the atlas frame width in terms of integer number of samples, where a sample corresponds to a luma sample of a video component. It is a requirement of V-PCC bitstream conformance that the value of asps\_frame\_width shall be equal to the value of vps\_frame\_width[ j ], where j is the index of the current atlas.

**asps\_frame\_height** indicates the atlas frame height in terms of integer number of samples, where a sample corresponds to a luma sample of a video component. It is a requirement of V-PCC bitstream conformance that the value of asps\_frame\_height shall be equal to the value of vps\_frame\_height[ j ], where j is the index of the current atlas.

During the reconstruction phase the decoded occupancy, geometry, and attribute videos shall be converted to the nominal width, height, and frame rate of the current atlas using appropriate scaling. Such scaling is outside the scope of this document.

**asps\_log2\_patch\_packing\_block\_size** specifies the value of the variable PatchPackingBlockSize, that is used for the horizontal and vertical placement of the patches within the atlas, as follows:

PatchPackingBlockSize = 2( asps\_log2\_patch\_packing\_block\_size) (7‑3)

The value of asps\_log2\_patch\_packing\_block\_size shall be in the range of 0 to 7, inclusive.

**asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4** specifies the value of the variable MaxAtlasFrmOrderCntLsb that is used in the decoding process for the atlas frame order count as follows:

MaxAtlasFrmOrderCntLsb = 2( asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4 + 4 ) (7‑4)

The value of asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4 shall be in the range of 0 to 12, inclusive.

**asps\_max\_dec\_atlas\_frame\_buffering\_minus1** plus 1 specifies the maximum required size of the decoded atlas frame buffer for the CAS in units of atlas frame storage buffers. The value of asps\_max\_dec\_atlas\_frame\_buffering\_minus1 shall be in the range of 0 to MaxDafbSize − 1, inclusive, where MaxDafbSize is as specified is Annex A.

**asps\_long\_term\_ref\_atlas\_frames\_flag** equal to 0 specifies that no long term reference atlas frame is used for inter prediction of any coded atlas frame in the CAS. asps\_long\_term\_ref\_atlas\_frames\_flagequal to 1 specifies that long term reference atlas frames may be used for inter prediction of one or more coded atlas frames in the CAS.

**asps\_num\_ref\_atlas\_frame\_lists\_in\_asps** specifies the number of the ref\_list\_struct( rlsIdx ) syntax structures included in the atlas sequence parameter set. The value of asps\_num\_ref\_atlas\_frame\_lists\_in\_asps shall be in the range of 0 to 64, inclusive.

NOTE – A decoder should allocate memory for a total number of ref\_list\_struct( rlsIdx ) syntax structures equal to (num\_ref\_atlas\_frame\_lists\_in\_asps + 1) since there may be one ref\_list\_struct(  rlsIdx ) syntax structure directly signalled in the atlas tile group headers of the current atlas tile group.

**asps\_use\_eight\_orientations\_flag** equal to 0 specifies that the patch orientation index for a patch with index j in a frame with index i, pdu\_orientation\_index[ i ][ j ], is in the range of 0 to 1, inclusive. asps\_use\_eight\_orientations\_flagequal to 1 specifies that the patch orientation index for a patch with index j in a frame with index i, pdu\_orientation\_index[ i ][ j ], is in the range of 0 to 7, inclusive.

**asps\_extended\_projection\_enabled\_flag** equal to 0 specifies that the patch projection information is not signalled for the current atlas tile group. asps\_extended\_projection\_enabled\_flag equal to 1 specifies that the patch projection information is signalled for the current atlas tile group. When asps\_extended\_projection\_enabled\_flag is not present, its value is inferred to be equal to 0.

**asps\_normal\_axis\_limits\_quantization\_enabled\_flag** equal to 1 specifies that quantization parameters shall be signalled and used for quantizing the normal axis related elements of a patch data unit, a merge patch data unit, or an inter patch data unit. If asps\_normal\_axis\_limits\_quantization\_enabled\_flag is equal to 0, then no quantization is applied on any normal axis related elements of a patch data unit, a merge patch data unit, or an inter patch data unit.

**asps\_normal\_axis\_max\_delta\_value\_enabled\_flag** equal to 1 specifies that the maximum nominal shift value of the normal axis that may be present in the geometry information of a patch with index i in a frame with index j will be indicated in the bitstream for each patch data unit, a merge patch data unit, or an inter patch data unit. If asps\_normal\_axis\_max\_delta\_value\_enabled\_flagis equal to 0 then the maximum nominal shift value of the normal axis that may be present in the geometry information of a patch with index i in a frame with index j shall not be be indicated in the bitstream for each patch data unit, a merge patch data unit, or an inter patch data unit.

**asps\_remove\_duplicate\_point\_enabled\_flag** equal to 1 indicates that duplicated points shall not be reconstructed for the current atlas, where a duplicated point is a point with the same 2D and 3D geometry coordinates as another point from a lower index map. asps\_remove\_duplicate\_point\_enabled\_flag equal to 0 indicates that all points shall be reconstructed.

[Ed. (AMT): This seems a "reconstruction" aspect that could have been handled using an SEI. Is it appropriate to have this signaled in the ASPS?]

[Ed. (AMT): It was agreed to remove this and move it in the geometry SEI. TBD]

[Ed. (AMT): regardless of moving this elsewhere, does this only apply to points with both the same 2D and 3D coordinates (so in correspondence to maps)? So what happens to when the same point exists in multiple patches? Duplication may even happen with data present in the same map. So restricting the statement to talk about "lower" index map might not be comprehensive enough. Clarify].

**asps\_pixel\_deinterleaving\_enabled\_flag** equal to 1 indicates that the decoded geometry and attribute videos for the current atlas contain spatially interleaved pixels from two maps. asps\_pixel\_deinterleaving\_flag equal to 0 indicates that the decoded geometry and attribute videos corresponding to the current atlas contain pixels from only a single map.

[Ed. (AMT): What is the interaction of this with vps\_map\_count\_minus1, vps\_multiple\_map\_streams\_present\_flag, and vps\_map\_absolute\_coding\_enabled\_flag? Should this not be present when vps\_map\_count\_minus1 is 0? What about vps\_map\_count\_minus1> 1? Would we then divide by two (in which case only even number of maps should be used)? Should also all maps be absolute encoded in this case? We likely would need to move this higher btw to deal with the interactions more elegantly]

[Ed. (RJ): This should be thought of from a decoder's perspective. So if asps\_pixel\_interleaving\_flag is 1 and decoder receives 2 maps, then each map has interleaved pixels from two maps. So on the encoder side, initially there would have been 4 maps. Please note that the Gothenburg decision was to have generalized text but have a restriction in the profiles so that when asps\_pixel\_interleaving\_flag is 1, there is only a single layer present.]

[Ed. (AMT): Should this explicitly mention "corresponding to a single video frame" instead of single stream? Note also that even the 0 case definition seems problematic since it talks about a single stream and a single map. This basically cancels the definition of vps\_multiple\_map\_streams\_present\_flag = 0]

[Ed. (ATA): Currently, CTC & the SW does not support more than 2 maps]

[Ed. (RJ): Currently the flag makes sense only when vps\_map\_count\_minus1 is equal to 0. In that case, the semantics seem ok. According to the Gothenburg decision, the text should be generalized to arbitrary number of maps and the restriction to two layers should be in profiles. In such a case, similar to the PLR case, there would be one flag indicating whether the geometry and attribute videos corresponding to each map is actually an interleaving of two maps. But this needs a new contribution.][Ed. (RJ): This was adopted in the Geneva October 2019 meeting.]

**asps\_patch\_precedence\_order\_flag** equal to 1 indicates that patch precedence for the current atlas is the same as the decoding order. asps\_patch\_precedence\_order\_flag equal to 0 indicates that patch precedence for the current atlas is the reverse of the decoding order.

**asps\_patch\_size\_quantizer\_present\_flag** equal to 1 indicates that the patch size quantization parameters are present in an atlas tile group header. If asps\_patch\_size\_quantizer\_present\_flag is equal to 0, then the patch size quantization parameters are not present.

**asps\_eom\_patch\_enabled\_flag**  equal to 1 indicates that the decoded occupancy map video for the current atlas contains information related to whether intermediate depth positions between two depth maps are occupied. asps\_eom\_patch\_enabled\_flag equal to 0 indicates that the decoded occupancy map video does not contain information related to whether intermediate depth positions between two depth maps are occupied.

It is a requirement of bitstream conformance that if asps\_eom\_patch\_enabled\_flag is equal to 1, oi\_lossy\_occupancy\_map\_compression\_threshold shall be equal to 0.

**asps\_point\_local\_reconstruction\_enabled\_flag** equal to 1 indicates that point local reconstruction mode information may be present in the bitstream for the current atlas. asps\_point\_local\_reconstruction\_enabled\_flag equal to 0 indicates that no information related to the point local reconstruction mode is present in the bitstream for the current atlas.

[Ed. (AMT): "may be used" seems more appropriate since it appears that the mode could be disabled if the interpolate\_flag and minimum\_depth are both set to 0 Ed. (PA): Changed as proposed but edit semantics on information presence in the bitstream. Usage or not is defined in profile sections.]

[Ed. (RJ): Need bitstream restrictions corresponding to values of asps\_pixel\_deinterleaving\_flag, asps\_point\_local\_reconstruction\_enabled\_flag, and asps\_enhanced\_occupancy\_map\_for\_depth\_enabled\_flag. ][Ed. (PA): Agreed, I think this could be done in the profile section when it applies -- like it was for wavefront, slices, tiles restrictions in HEVC although the intent was a bit different]

[Ed. (RJ): Added bitstream conformance condition. If the specification works for a combination of flags then it may be appropriate to have the restriction in the profiles. Otherwise it should be in semantics.]

It is a requirement of bitstream conformance that asps\_pixel\_deinterleaving\_flag and asps\_point\_local\_reconstruction\_enabled\_flag shall not be one simultaneously.

**asps\_map\_count\_minus1** plus 1 indicates the number of maps that may be used for encoding the geometry and attribute data for the current atlas. asps\_map\_count\_minus1 shall be in the range of 0 to 15, inclusive. It is a requirement of bitstream conformance to this document that asps\_map\_count\_minus1 is equal to vps\_map\_count\_minus1[ atlasId ], where atlasId is the index of the current atlas.

**asps\_pixel\_deinterleaving\_map\_flag**[ i ] equal to 1 indicates that decoded geometry and attribute videos correpsonding to map with index i in the current atlas contain spatially interleaved pixels corresponding to two maps. asps\_pixel\_deinterleaving\_map\_flag[ i ] equal to 0 indicates that decoded geometry and attribute videos corresponding to map index i in the current atlas contain pixels corresponding to a single map. When not present, the value of asps\_pixel\_deinterleaving\_map\_flag[ i ] is inferred to be 0.

**asps\_eom\_fix\_bit\_count\_minus1** plus 1 indicates the size in bits of the EOM codeword. asps\_enhanced\_occupancy\_map\_fix\_bit\_count shall be in the range of 0 to Min(15, oi\_occupancy\_nominal\_2d\_bitdepth\_minus1[ atlasId ] − 1), inclusive, where atlasId is the index of the current atlas.

[Ed. (PA): Check range.] [Ed. (PA): Range fixed, this syntax element is u(4)]

**asps\_surface\_thickness\_minus1** plus 1 specifies the maximum absolute difference between an explicitly coded depth value and interpolated depth value when asps\_pixel\_deinterleaving\_enabled\_flag or asps\_point\_local\_reconstruction\_enabled\_flag is equal to 1.

**asps\_vui\_parameters\_present\_flag** equal to 1 specifies that the vui\_parameters( ) syntax structure as specified in Annex G is present. asps\_vui\_parameters\_present\_flag equal to 0 specifies that the vui\_parameters( ) syntax structure as specified in Annex E is not present.

**asps\_extension\_flag** equal to 0 specifies that no asps\_extension\_data\_flag syntax elements are present in the ASPS RBSP syntax structure. asps\_extension\_flag shall be equal to 0 in bitstreams conforming to this version of this document. The value of 1 for asps\_extension\_\_present\_flag is reserved for future use by ISO/IEC. Decoders shall ignore all asps\_extension\_data\_flag syntax elements that follow the value 1 for asps\_extension\_flag in an ASPS NAL unit.

**asps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in this version of this document. Decoders conforming to this version of this document shall ignore all asps\_extension\_data\_flag syntax elements.

##### Point local reconstruction information semantics

**plri\_point\_local\_reconstruction\_map\_flag**[ i ] equal to 1 indicates that point local reconstruction mode information is present in the bitstream for the map of index i of the current atlas. plri\_point\_local\_reconstruction\_map\_flag equal to 0 indicates that no information related to the point local reconstruction mode is present in the bitstream for the map of index i of the current atlas.

[Ed. (RJ): My suggestion is that "enabled" be removed from this name. That way, you have an enabling flag at the asps level and normal flag at the map level. I have made similar change for pixel deinterleaving.]

[Ed. (PA): OK- proceeded as proposed.]

**plri\_number\_of\_modes\_minus1**[ mapIdx ] plus 1 indicates the number of reconstruction modes specified for the point local reconstruction process associated with the map of index mapIdx of the current atlas. The value of plri\_number\_of\_modes\_minus1[ mapIdx ] shall be in the range of 0 to 15, inclusive.

**plri\_interpolate\_flag**[ mapIdx ][ j ]equal to 1 indicates that point interpolation for the map of index mapIdx of the current atlas may be used for point local reconstruction mode j. plri\_interpolate\_flag[ mapIdx ][ j ]equal to 0 indicates that no point interpolation for the map of index mapIdx of the current atlas is used for point local reconstruction mode j.

**plri\_filling\_flag**[ mapIdx ][ j ] equal to 1 indicates that filling for the map of index mapIdx of the current atlas may be used for point local reconstruction mode j. plri\_filling\_flag[ mapIdx ][ j ] equal to 0 indicates that no filling for the map of index mapIdx of the current atlas is used for point local reconstruction mode j.

**plri\_minimum\_depth**[ mapIdx ][ j ] specifies the value of the minimum depth parameter of the point local reconstruction mode j for the map of index mapIdx of the current atlas. plri\_minimum\_depth[ mapIdx ][ j ] shall be in the range of 0 to 3, inclusive.

**plri\_neighbour\_minus1**[ mapIdx ][ j ] plus 1 specifies the size of the 2D neighbourhood for point local construction mode j for the map of index mapIdx of the current atlas. plri\_neighbour\_minus1[ mapIdx ][ j ] shall be in the range of 0 to 3, inclusive.

**plri\_block\_threshold\_per\_patch\_minus1**[ mapIdx ] plus 1 specifies the value representative of the threshold used to define the values contained in the array plrd\_level. plri\_block\_threshold\_per\_patch\_minus1[ mapIdx ] shall be in the range of 0 to 63, inclusive.

#### Atlas frame parameter set RBSP semantics

##### General atlas frame parameter set RBSP semantics

**afps\_atlas\_frame\_parameter\_set\_id** identifies the atlas frame parameter set for reference by other syntax elements. The value of afps\_atlas\_frame\_parameter\_set\_idshall be in the range of 0 to 63, inclusive.

**afps\_atlas\_sequence\_parameter\_set\_id** specifies the value of asps\_atlas\_sequence\_parameter\_set\_id for the active atlas sequence parameter set. The value of afps\_atlas\_sequence\_parameter\_set\_id shall be in the range of 0 to 15, inclusive.

**afps\_output\_flag\_present\_flag** equal to 1 indicates that the atgh\_frame\_output\_flag syntax element is present in the associated tile group headers. afps\_output\_flag\_present\_flag equal to 0 indicates that the atgh\_frame\_output\_flag syntax element is not present in the associated tile group headers.

**afps\_num\_ref\_idx\_default\_active\_minus1** plus 1 specifies the inferred value of the variable NumRefIdxActive for the tile group with atgh\_num\_ref\_idx\_active\_override\_flag equal to 0. The value of afps\_num\_ref\_idx\_default\_active\_minus1 shall be in the range of 0 to 14, inclusive.

**afps\_additional\_lt\_afoc\_lsb\_len** specifies the value of the variable MaxLtAtlasFrmOrderCntLsb that is used in the decoding process for reference atlas frame lists as follows:

MaxLtAtlasFrmOrderCntLsb = 2( asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4 + 4 + afps\_additional\_lt\_afoc\_lsb\_len) (7‑5)

The value of afps\_additional\_lt\_afoc\_lsb\_len shall be in the range of 0 to 32 – asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4 − 4, inclusive. When asps\_long\_term\_ref\_atlas\_frames\_flag is equal to 0, the value of afps\_additional\_lt\_afoc\_lsb\_len shall be equal to 0.

**afps\_3d\_pos\_x\_bit\_count\_minus1** plus 1 specifies the number of bits in the fixed-length representation of pdu\_3d\_pos\_x[ j ] of patch with index j in an atlas tile group that refers to afps\_atlas\_frame\_parameter\_set\_id. The value of afps\_3d\_pos\_x\_bit\_count\_minus1 shall be in the range of 0 to 31, inclusive.

**afps\_3d\_pos\_y\_bit\_count\_minus1** plus 1 specifies the number of bits in the fixed-length representation of pdu\_3d\_pos\_y[ j ] of patch with index j in an atlas tile group that refers to afps\_atlas\_frame\_parameter\_set\_id. The value of afps\_3d\_pos\_y\_bit\_count\_minus1 shall be in the range of 0 to 31, inclusive.

**afps\_lod\_mode\_enabled\_flag** equal to 1 indicates that the LOD parameters may be present in a patch. afps\_lod\_mode\_enabled\_flag equal to 0 indicates that the LOD parameters shall not be present in a patch.

**afps\_override\_eom\_for\_depth\_flag** equal to 1 indicates that the values of afps\_eom\_number\_of\_patch\_bit\_count\_minus1 and afps\_eom\_max\_bit\_count\_minus1 is explicitly present in the bitstream. afps\_override\_eom\_for\_depth\_flag equal to 0 indicates that the values of afps\_eom\_number\_of\_patch\_bit\_count\_minus1 and afps\_eom\_max\_bit\_count\_minus1 are implicitly derived.

**afps\_eom\_number\_of\_patch\_bit\_count\_minus1** plus 1 specifies the number of bits used to represent the number of geometry patches associated in an EOM attribute patch in an atlas frame that is associated with this atlas frame parameter set. afps\_eom\_number\_of\_patch\_bit\_count\_minus1 shall be in the range of 0 to 15, inclusive. When not present, the value of afps\_eom\_number\_of\_patch\_bit\_count\_minus1 is inferred to be equal to 7.

**afps\_eom\_max\_bit\_count\_minus1** plus 1 specifies the number of bits used to represent the number of EOM points per geometry patch associated with an EOM attribute patch in an atlas frame that is associated with this atlas frame parameter set. afps\_eom\_max\_bit\_count\_minus1 shall be in the range of 0 to 15, inclusive. When not present, the value of afps\_eom\_max\_bit\_count\_minus1 is inferred to be equal to 7.

**afps\_raw\_3d\_pos\_bit\_count\_explicit\_mode\_flag** equal to 1 indicates that the number of bits in the fixed-length representation of rpdu\_3d\_pos\_x, rpdu\_3d\_pos\_y, and rpdu\_3d\_pos\_z is explicitly coded by atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 in the atlas tile group header that refers to afps\_atlas\_frame\_parameter\_set\_id. afps\_raw\_3d\_pos\_bit\_count\_explicit\_mode\_flag equal to 0 indicates the value of atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 is implicitly derived.

**afps\_extension\_flag** equal to 0 specifies that no afps\_extension\_data\_flag syntax elements are present in the AFPS RBSP syntax structure. afps\_extension\_flag shall be equal to 0 in bitstreams conforming to this version of this document. The value of 1 for afps\_extension\_present\_flag is reserved for future use by ISO/IEC. Decoders shall ignore all afps\_extension\_data\_flag syntax elements that follow the value 1 for afps\_extension\_flag in an AFPS NAL unit.

**afps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in this version of this document. Decoders conforming to this version of this document shall ignore all afps\_extension\_data\_flag syntax elements.

##### Atlas frame tile information syntax

**afti\_single\_tile\_in\_atlas\_frame\_flag** equal to 1 specifies that there is only one tile in each atlas frame referring to the AFPS. afti\_single\_tile\_in\_atlas\_frame\_flag equal to 0 specifies that there is more than one tile in each atlas frame referring to the AFPS.

**afti\_uniform\_tile\_spacing\_flag** equal to 1 specifies that tile column and row boundaries are distributed uniformly across the atlas frame and signalled using the syntax elements afti\_tile\_cols\_width\_minus1 and afti\_tile\_rows\_height\_minus1, respectively. afti\_uniform\_tile\_spacing\_flag equal to 0 specifies that tile column and row boundaries may or may not be distributed uniformly across the atlas frame and are signalled using the syntax elements afti\_num\_tile\_columns\_minus1 and afti\_num\_tile\_rows\_minus1 and a list of syntax element pairs afti\_tile\_column\_width\_minus1[ i ] and afti\_tile\_row\_height\_minus1[ i ]. When not present, the value of afti\_uniform\_tile\_spacing\_flag is inferred to be equal to 1.

**afti\_tile\_cols\_width\_minus1** plus 1 specifies the width of the tile columns excluding the right-most tile column of the atlas frame in units of 64 samples when afti\_uniform\_tile\_spacing\_flag is equal to 1. The value of afti\_tile\_cols\_width\_minus1 shall be in the range of 0 to ( asps\_frame\_width + 63 ) / 64 − 1, inclusive. When not present, the value of afti\_tile\_cols\_width\_minus1 is inferred to be equal to ( asps\_frame\_width + 63 ) / 64 – 1**.**

**afti\_tile\_rows\_height\_minus1** plus 1 specifies the height of the tile rows excluding the bottom tile row of the atlas frame in units of 64 samples when afti\_uniform\_tile\_spacing\_flag is equal to 1. The value of afti\_tile\_rows\_height\_minus1 shall be in the range of 0 to ( asps\_frame\_height + 63 ) / 64 − 1, inclusive. When not present, the value of afti\_tile\_rows\_height\_minus1 is inferred to be equal to ( asps\_frame\_height + 63 ) / 64 – 1.

**afti\_num\_tile\_columns\_minus1** plus 1 specifies the number of tile columns partitioning the atlas frame when afti\_uniform\_tile\_spacing\_flag is equal to 0. The value of afti\_num\_tile\_columns\_minus1 shall be in the range of 0 to ( asps\_frame\_width + 63 ) / 64 − 1, inclusive. If afti\_single\_tile\_in\_atlas\_frame\_flag is equal to 1, the value of afti\_num\_tile\_columns\_minus1 is inferred to be equal to 0. Otherwise, when afti\_uniform\_tile\_spacing\_flag is equal to 1, the value of afti\_num\_tile\_columns\_minus1 is inferred as specified in clause 6.4.1.

**afti\_num\_tile\_rows\_minus1** plus 1 specifies the number of tile rows partitioning the atlas frame when pti\_uniform\_tile\_spacing\_flag is equal to 0. The value of pti\_num\_tile\_rows\_minus1 shall be in the range of 0 to ( asps\_frame\_height + 63 ) / 64 − 1, inclusive. If afti\_single\_tile\_in\_pic\_flag is equal to 1, the value of afti\_num\_tile\_rows\_minus1 is inferred to be equal to 0. Otherwise, when afti\_uniform\_tile\_spacing\_flag is equal to 1, the value of afti\_num\_tile\_rows\_minus1 is inferred as specified in clause 6.4.1.

The variable NumTilesInAtlasFrame is set equal to ( afti\_num\_tile\_columns\_minus1 + 1 ) \* ( afti\_num\_tile\_rows\_minus1 + 1 ).

When afti\_single\_tile\_in\_atlas\_frame\_flag is equal to 0, NumTilesInAtlasFrame shall be greater than 1.

**afti\_tile\_column\_width\_minus1**[ i ] plus 1 specifies the width of the i-th tile column in units of 64 samples.

**afti\_tile\_row\_height\_minus1**[ i ] plus 1 specifies the height of the i-th tile row in units of 64 samples.

**afti\_single\_tile\_per\_tile\_group\_flag** equal to 1 specifies that each tile group that refers to this AFPS includes one tile. afti\_single\_tile\_per\_tile\_group\_flag equal to 0 specifies that a tile group that refers to this AFPS may include more than one tile. When not present, the value of afti\_single\_tile\_per\_tile\_group\_flag is inferred to be equal to 1.

**afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1** plus 1 specifies the number of tile groups in each atlas frame referring to the AFPS. The value of afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1shall be in the range of 0 to NumTilesInAtlasFrame − 1, inclusive. When not present and afti\_single\_tile\_per\_tile\_group\_flag is equal to 1, the value of afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1is inferred to be equal to NumTilesInAtlasFrame − 1.

**afti\_top\_left\_tile\_idx**[ i ] specifies the tile index of the tile located at the top-left corner of the i-th tile group. The value of afti\_top\_left\_tile\_idx[ i ] shall not be equal to the value of afti\_top\_left\_tile\_idx[ j ] for any i not equal to j. When not present, the value of afti\_top\_left\_tile\_idx[ i ] is inferred to be equal to i. The length of the afti\_top\_left\_tile\_idx[ i ] syntax element is Ceil( Log2( NumTilesInAtlasFrame ) bits.

**afti\_bottom\_right\_tile\_idx\_delta**[ i ] specifies the difference between the tile index of the tile located at the bottom-right corner of the i-th tile group and afti\_top\_left\_tile\_idx[ i ]. When afti\_single\_tile\_per\_tile\_group\_flag is equal to 1, the value of afti\_bottom\_right\_tile\_idx\_delta[ i ] is inferred to be equal to 0. The length of the afti\_bottom\_right\_tile\_idx\_delta[ i ] syntax element is Ceil( Log2( NumTilesInAtlasFrame − afti\_top\_left\_tile\_idx[ i ] ) ) bits.

The variables TopLeftTileColumn[ i ], TopLeftTileRow[ i ], BottomRightTileColumn[ i ], and BottomRightTileRow[ i ], which specify the corresponding tile column and row positions for the top left an bottom right tiles in a tile group are computed as follows:

TopLeftTileColumn[ i ] = afti\_top\_left\_tile\_idx[ i ] % (afti\_num\_tile\_rows\_minus1 + 1)  
 TopLeftTileRow[ i ] = afti\_top\_left\_tile\_idx[ i ] / (afti\_num\_tile\_rows\_minus1  + 1)   
 botRightTileIdx = afti\_top\_left\_tile\_idx[ i ] + afti\_bottom\_right\_tile\_idx\_delta[ i ]  
 BottomRightTileColumn[ i ] = botRightTileIdx % (afti\_num\_tile\_rows\_minus1 + 1)  
 BottomRightTileRow[ i ] = botRightTileIdx / (afti\_num\_tile\_rows\_minus1  + 1)

The variables NumTilesInTileGroups[ i ] and TilesToTileGroupMap[ j ], which specify the number of tiles in the i-th tile group and the mapping of tiles to tile groups, are derived as follows:

NumTilesInTileGroups[ i ] = 0  
 for( j = 0; j < NumTilesInAtlasFrame; j++) {  
 if( TileColBd[ j ] >= TileColBd [ afti\_top\_left\_tile\_idx[ i ] ] &&  
 TileColBd [ j ] <= TileColBd [ botRightTileIdx] &&  
 TileRowBd[ j ] >= TileRowBd [ afti\_top\_left\_tile\_idx[ i ] ] && (7‑6)  
 TileRowBd [ j ] <= TileRowBd [ botRightTileIdx] ) {  
 NumTilesInTileGroups[ i ]++  
 TilesToTileGroupMap[ j ] = i  
 }  
 }

The variables TileGroupWidth[ i ] and TileGroupHeight[ i ], which specify the width and height of a tile group respectively, are then computed as follows:

TileGroupWidth[ i ] = 0   
 TileGroupHeight[ i ] = 0  
 for( j= TopLeftTileColumn[ i ]; j <= BottomRightTileColumn[ i ]; j++) {  
 TileGroupWidth[ i ] += ColWidth[ j ] \* 64  
 }  
 for( j= TopLeftTileRow[ i ]; j <= BottomRightTileRow[ i ]; j++) {  
 TileGroupHeight[ i ] += RowHeight[ j ] \* 64  
 }

**afti\_signalled\_tile\_group\_id\_flag** equal to 1 specifies that the tile group ID for each tile group is signalled. afti\_signalled\_tile\_group\_id\_flag equal to 0 specifies that tile group IDs are not signalled.

**afti\_signalled\_tile\_group\_id\_length\_minus1** plus 1 specifies the number of bits used to represent the syntax element afti\_tile\_group\_id[ i ] when present, and the syntax element atgh\_address in tile group headers. The value of afti\_signalled\_tile\_group\_id\_length\_minus1 shall be in the range of 0 to 15, inclusive. When not present, the value of afti\_signalled\_tile\_group\_id\_length\_minus1 is inferred to be equal to Ceil( Log2( afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1 + 1 ) ) − 1.

**afti\_tile\_group\_id**[ i ] specifies the tile group ID of the i-th tile group. The length of the afti\_tile\_group\_id[ i ] syntax element is afti\_signalled\_tile\_group\_id\_length\_minus1 + 1 bits. When not present, the value of afti\_tile\_group\_id[ i ] is inferred to be equal to i, for each i in the range of 0 to afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1, inclusive.

#### Atlas adaptation parameter set RBSP semantics

##### General atlas adaptation parameter set RBSP semantics

**aaps\_atlas\_adaptation\_parameter\_set\_id** identifies the atlas adaptation parameter set for reference by other syntax elements. The value of acps\_atlas\_adaptation\_parameter\_set\_idshall be in the range of 0 to 63, inclusive.

**~~aaps\_atlas\_sequence\_parameter\_set\_id~~** ~~specifies the value of asps\_atlas\_sequence\_parameter\_set\_id for the active atlas sequence parameter set. The value of aaps\_atlas\_sequence\_parameter\_set\_id shall be in the range of 0 to 15, inclusive.~~

[Ed. (AMT): Need to discuss if we should keep the SPS id here or not. MIV currently does not have it. Could we have "orphaned" parameters? Do we want to allow parameters to be used across atlases/sps without resignaling these? If so, then we can remove this. I have it crossed and highlighted since my assumption is that this is to be removed]

**aaps\_camera\_parameters\_present\_flag** equal to 1 specifies that camera parameters shall be present in the current atlas adaptation parameter set. aaps\_camera\_parameters\_present\_flag equal to 0 specifies that camera parameters for the currrent adaptation parameter set shall not be present.

**aaps\_extension\_flag** equal to 0 specifies that no aaps\_extension\_data\_flag syntax elements are present in the AAPS RBSP syntax structure. aaps\_extension\_flag shall be equal to 0 in bitstreams conforming to this version of this document. The value of 1 for aaps\_extension\_present\_flag is reserved for future use by ISO/IEC. Decoders shall ignore all aaps\_extension\_data\_flag syntax elements that follow the value 1 for aaps\_extension\_flag in an AAPS NAL unit.

**aaps\_extension\_data\_flag** may have any value. Its presence and value do not affect decoder conformance to profiles specified in this version of this document. Decoders conforming to this version of this document shall ignore all aaps\_extension\_data\_flag syntax elements.

##### Atlas camera parameters semantics

**acp\_camera\_model** indicates the camera model for point cloud frames that are associated with the current adaptation parameter set as listed in Table 7‑4**.** The value of acp\_camera\_model shall be equal to 0 or 1 in bitstreams conforming to this version of this document. Other values of acp\_camera\_model are reserved for future use by ISO/IEC. Decoders conforming to this version of this document shall ignore reserved values of acp\_camera\_model.

Table 7‑4 – Interpretation of acp\_camera\_model

|  |  |
| --- | --- |
| **acp\_camera\_model** | **Name of acp\_camera\_model** |
| 0 | UNSPECIFIED |
| 1 | Orthographic camera model |
| 2 - 255 | RESERVED |

**acp\_scale\_enabled\_flag** equal to 1 indicates that scale parameters for the current camera model are present. acp\_scale\_enabled\_flagequal to 0 indicates that scale parameters for the current camera model are not present. When acp\_scale\_enabled\_flagis not present, it shall be inferred to be equal to 0.

**acp\_offset\_enabled\_flag** equal to 1 indicates that offset parameters for the current camera model are present. acp\_offset\_enabled\_flagequal to 0 indicates that offset parameters for the current camera model are not present. When acp\_offset\_enabled\_flagis not present, it shall be inferred to be equal to 0.

**acp\_rotation\_enabled\_flag** equal to 1 indicates that rotation parameters for the current camera model are present. acp\_rotation\_enabled\_flagequal to 0 indicates that rotation parameters for the current camera model are not present. When acp\_rotation\_enabled\_flagis not present, it shall be inferred to be equal to 0.

**acp\_scale\_on\_axis**[ d ] specifies the value of the scale, Scale[ d ], along the d axis for the current camera model in increments of 1 ÷ 216. The value of d is in the range of 0 to 2, inclusive, with the values of 0, 1, and 2 corresponding to the X, Y, and Z axis, respectively.The value of acp\_scale\_on\_axis[ d ], shall be in the range of 1 to 232 – 1, inclusive, where When acp\_scale\_on\_axis[ d ] is not present, it shall be inferred to be equal to 216. The value of Scale[ d ] is computed as follows:

Scale[ d ] = acp\_scale\_on\_axis[ d ] ÷ 216

A ScaleMatrix can be represented as follows:

**acp\_offset\_on\_axis**[ d ]indicates the value of the offset, Offset[d], along the d axis for the current camera model. The value of acp\_offset\_on\_axis[ d ] shall be in the range of −231 to 231 − 1, inclusive, where d is in the range of 0 to 2, inclusive. The values of d equal to 0, 1, and 2 correspond to the X, Y, and Z axis, respectively. When acp\_offset\_on\_axis[ d ] is not present, it shall be inferred to be equal to 0.

An offset matrix can be represented as follows:

**acp\_rotation\_qx** specifies the x component, qX, for the rotation of the current camera model using the quaternion representation. The value of acp\_rotation\_qx shall be in the range of −215 to 215 − 1, inclusive. When acp\_rotation\_qx is not present, its value shall be inferred to be equal to 0. The value of qX is computed as follows:

qX = acp\_rotation\_qx ÷ 215

**acp\_rotation\_qy** specifies the y component, qY, for the rotation of the current camera model using the quaternion representation. The value of acp\_rotation\_qy shall be in the range of −215 to 215 − 1, inclusive. When acp\_rotation\_qy is not present, its value shall be inferred to be equal to 0. The value of qY is computed as follows:

qY = acp\_rotation\_qy ÷ 215

**acp\_rotation\_qz** specifies the z component, qZ, for the rotation of the current camera model using the quaternion representation. The value of acp\_rotation\_qz shall be in the range of −215 to 215 − 1, inclusive. When acp\_rotation\_qz is not present, its value shall be inferred to be equal to 0. The value of qZ is computed as follows:

qZ = acp\_rotation\_qz ÷ 215

The fourth component, qW, for the rotation of the current camera model using the quaternion representation is calculated as follows:

qW = Sqrt( 1 – ( qX2 + qY2 + qZ2 ) )

A unit quaternion can be represented as a rotation matrix R as follows:

When an input point, OriginalPoint, is represented as a 4x1 column vector [x, y, z, 1], then the output point, TransformPoint, can be calculated as follows:

TransformedPoint = OffsetMatrix \* RotationMatrix \* ScaleMatrix \* OriginalPoint

[Ed. Note: The application of the Offset, Rotation, and ScaleMatrix should likely happen in Annex B]

#### Supplemental enhancement information RBSP semantics

Supplemental enhancement information (SEI) contains information that is not necessary to decode the samples of coded atlas frames from ACL NAL units. An SEI RBSP contains one or more SEI messages.

#### Access unit delimiter RBSP semantics

The access unit delimiter may be used to indicate the type of tile groups present in the coded atlas frames in the access unit containing the access or V-PCC access unit delimiter NAL unit and to simplify the detection of the boundary between access units. There is no normative decoding process associated with the access unit delimiter.

**aframe\_type** indicates that the atgh\_type values for all tile groups of the coded atlas frame in the access unit containing the access unit delimiter or V-PCC access unit delimiter NAL unit are members of the set listed in Table 7‑4 for the given value of aframe\_type. The value of aframe\_type shall be equal to 0, 1, 2, or 3 in bitstreams conforming to this version of this document. Other values of aframe\_type are reserved for future use by ISO/IEC. Decoders conforming to this version of this document shall ignore reserved values of aframe\_type.

Table 7‑4 – Interpretation of aframe\_type

|  |  |
| --- | --- |
| **aframe\_type** | **atgh\_type values that may be present in the coded atlas frame** |
| 0 | I\_TILE\_GRP |
| 1 | P\_TILE\_GRP, I\_TILE\_GRP |
| 2 | SKIP\_TILE\_GRP, P\_TILE\_GRP, I\_TILE\_GRP |
| 3 | SKIP\_TILE\_GRP |
| 4 - 7 | RESERVED |

#### End of sequence RBSP semantics

When present, the end of sequence RBSP specifies that the current access unit is the last access unit in the coded atlas sequence in decoding order and the next subsequent access unit in the bitstream in decoding order (if any) is an IRAP access unit. The syntax content of the SODB and RBSP for the end of sequence RBSP are empty.

#### End of bitstream RBSP semantics

The end of bitstream RBSP indicates that no additional NAL units are present in the bitstream that are subsequent to the end of bitstream RBSP in decoding order. The syntax content of the SODB and RBSP for the end of bitstream RBSP are empty.

#### Filler data RBSP semantics

The filler data RBSP contains bytes whose value shall be equal to 0xFF. No normative decoding process is specified for a filler data RBSP.

**ff\_byte** is a byte equal to 0xFF.

#### Atlas tile group layer rbsp semantics

None.

#### Atlas tile group header semantics

**atgh\_atlas\_frame\_parameter\_set\_id** specifies the value of afps\_atlas\_frame\_parameter\_set\_id for the active atlas frame parameter set for the current atlas tile group. The value of atgh\_atlas\_frame\_parameter\_set\_id shall be in the range of 0 to 63, inclusive.

**atgh\_atlas\_adaptation\_parameter\_set\_id** specifies the value of aaps\_atlas\_adaptation\_parameter\_set\_id for the active atlas adaptation parameter set for the current atlas tile group. The value of atgh\_atlas\_adaptation\_parameter\_set\_idshall be in the range of 0 to 63, inclusive. It is a requirement of bitstream conformance that all tile groups of an atlas frame shall have the same atgh\_atlas\_adaptation\_parameter\_set\_id.

[Ed. (AMT): We should likely disallow permitting different values so as to not make things complicated with aspects such as smoothing. There is also the question if this should be included as part of conformance. TBD]

**atgh\_address** specifies the tile group address of the tile group. When not present, the value of atgh\_addressis inferred to be equal to 0.

The following applies:

* The tile group address is the tile group ID of the tile group.
* The length of atgh\_address is afti\_signalled\_tile\_group\_id\_length\_minus1 + 1 bits.
* If afti\_signalled\_tile\_group\_id\_flag is equal to 0, the value of atgh\_address shall be in the range of 0 to afti\_num\_tile\_groups\_in\_atlas\_frame\_minus1, inclusive. Otherwise, the value of atgh\_address shall be in the range of 0 to 2( afti\_signalled\_tile\_group\_id\_length\_minus1 + 1 ) − 1, inclusive.

It is a requirement of bitstream conformance that the following constraints apply:

* The value of atgh\_address shall not be equal to the value of atgh\_address of any other coded atlas tile group unit of the same coded atlas frame.
* The tile groups of an atlas frame shall be in increasing order of their atgh\_address values.

**atgh\_type** specifies the coding type of the current atlas tile goup according to Table 7‑5. The value of atgh\_type shall be equal to 0, 1, or 2 in bitstreams conforming to this version of this document. Other values of atgh\_type are reserved for future use by ISO/IEC. Decoders conforming to this version of this document shall ignore reserved values of atgh\_type.

Table 7‑5 – Name association to atgh\_type

|  |  |
| --- | --- |
| atgh\_type | Name of atgh\_type |
| 0 | P\_TILE\_GRP (Inter atlas tile group) |
| 1 | I\_TILE\_GRP (Intra atlas tile group) |
| 2 | SKIP\_TILE\_GRP (SKIP atlas tile group) |
| 3 - … | RESERVED |

**atgh\_atlas\_output\_flag** affects the decoded atlas output and removal processes as specified in Annex E. When atgh\_atlas\_output\_flag is not present, it is inferred to be equal to 1.

**atgh\_atlas\_frm\_order\_cnt\_lsb** specifies the atlas frame order count modulo MaxAtlasFrmOrderCntLsb for the current atlas tile group. The length of the atgh\_atlas\_frm\_order\_cnt\_lsb syntax element is equal to asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4 + 4 bits. The value of the atgh\_atlas\_frm\_order\_cnt\_lsb shall be in the range of 0 to MaxAtlasFrmOrderCntLsb − 1, inclusive. When atgh\_atlas\_frm\_order\_cnt\_lsb is not present, it shall be inferred to be equal to 0.

**atgh\_ref\_atlas\_frame\_list\_sps\_flag** equal to 1 specifies that the reference atlas frame list of the current atlas tile group is derived based on one of the ref\_list\_struct( rlsIdx ) syntax structures in the active ASPS. atgh\_ref\_atlas\_frame\_list\_sps\_flag equal to 0 specifies that the reference atlas frame list of the current atlas tile list is derived based on the ref\_list\_struct( rlsIdx ) syntax structure that is directly included in the tile group header of the current atlas tile group. When asps\_num\_ref\_atlas\_frame\_lists\_in\_aspsis equal to 0, the value of atgh\_ref\_atlas\_frame\_list\_sps\_flag is inferred to be equal to 0.

**atgh\_ref\_atlas\_frame\_list\_idx** specifies the index, into the list of the ref\_list\_struct( rlsIdx ) syntax structures included in the active ASPS, of the ref\_list\_struct( rlsIdx ) syntax structure that is used for derivation of the reference atlas frame list for the current atlas tile group. The syntax element atgh\_ref\_atlas\_frame\_list\_idx is represented by Ceil( Log2( asps\_num\_ref\_atlas\_frame\_lists\_in\_asps ) ) bits. When not present, the value of atgh\_ref\_atlas\_frame\_list\_idx is inferred to be equal to 0. The value of atgh\_ref\_atlas\_frame\_list\_idx shall be in the range of 0 to asps\_num\_ref\_atlas\_frame\_lists\_in\_asps − 1, inclusive. When atgh\_ref\_atlas\_frame\_list\_sps\_flag is equal to 1 and asps\_num\_ref\_atlas\_frame\_lists\_in\_asps is equal to 1, the value of atgh\_ref\_atlas\_frame\_list\_idx is inferred to be equal to 0.

The variable RlsIdx for the current atlas tile group is derived as follows:

RlsIdx = asps\_num\_ref\_atlas\_frame\_lists\_in\_asps ?   
 atgh\_ref\_atlas\_frame\_list\_idx : asps\_num\_ref\_atlas\_frame\_lists\_in\_asps (7‑7)

**atgh\_additional\_afoc\_lsb\_present\_flag**[ j ] equal to 1 specifies that atgh\_additional\_afoc\_lsb\_val[ j ] is present for the current atlas tile group. atgh\_additional\_afoc\_lsb\_present\_flag[ j ] equal to 0 specifies that atgh\_additional\_afoc\_lsb\_val[ j ] is not present.

**atgh\_additional\_afoc\_lsb\_val**[ j ] specifies the value of FullAtlasFrmOrderCntLsbLt[ RlsIdx ][ j ] for the current atlas tile group as follows:

FullAtlasFrmOrderCntLsbLt[ RlsIdx ][ j ] =   
 atgh\_additional\_afoc\_lsb\_val[ j ] \* MaxAtlasFrmOrderCntLsb +  
 afoc\_lsb\_lt[ RlsIdx ][ j ] (7‑8)

The syntax element atgh\_additional\_afoc\_lsb\_val[ j ] is represented by afps\_additional\_lt\_afoc\_lsb\_len bits. When not present, the value of atgh\_additional\_afoc\_lsb\_val[ j ] is inferred to be equal to 0.

**atgh\_pos\_min\_z\_quantizer** specifies the quantizer that is to be applied to the pdu\_3d\_pos\_min\_z[ p ] value of the patch p. If atgh\_pos\_min\_z\_quantizer is not present, its value shall be inferred to be equal to 0.

**atgh\_pos\_delta\_max\_z\_quantizer** specifies the quantizer that is to be applied to the pdu\_3d\_pos\_delta\_max\_z[ p ] value of the patch with index p. If atgh\_pos\_delta\_max\_z\_quantizer is not present, its value shall be inferred to be equal to 0.

**atgh\_patch\_size\_x\_info\_quantizer** specifies the value of the quantizer PatchSizeXQuantizer that is to be applied to the variables pdu\_2d\_size\_x\_minus1[ p ], mpdu\_2d\_delta\_size\_x[ p ], ipdu\_2d\_delta\_size\_x[ p ], rpdu\_2d\_size\_x\_minus1[ p ], and epdu\_2d\_size\_x\_minus1[ p ] of a patch with index p. If atgh\_patch\_size\_x\_info\_quantizer is not present, its value shall be inferred to be equal to asps\_log2\_patch\_packing\_block\_size. The value of PatchSizeXQuantizer is computed as follows:

PatchSizeXQuantizer = 1 << atgh\_patch\_size\_x\_info\_quantizer (7‑9)

The value of atgh\_patch\_size\_x\_info\_quantizer shall be in the range of 0 to asps\_log2\_patch\_packing\_block\_size, inclusive.

**atgh\_patch\_size\_y\_info\_quantizer** specifies the value of the quantizer PatchSizeYQuantizer that is to be applied to the variables pdu\_2d\_size\_y\_minus1[ p ], mpdu\_2d\_delta\_size\_y[ p ], ipdu\_2d\_delta\_size\_y[ p ], rpdu\_2d\_size\_y\_minus1[ p ], and epdu\_2d\_size\_y\_minus1[ p ] of a patch with index p. If atgh\_patch\_size\_y\_info\_quantizer is not present, its value shall be inferred to be equal to asps\_log2\_patch\_packing\_block\_size. The value of PatchSizeYQuantizer is computed as follows:

PatchSizeYQuantizer = 1 << atgh\_patch\_size\_y\_info\_quantizer (7‑10)

The value of atgh\_patch\_size\_y\_info\_quantizer shall be in the range of 0 to asps\_log2\_patch\_packing\_block\_size, inclusive.

**atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1** plus 1 specifies the number of bits in the fixed-length representation of rpdu\_3d\_pos\_x, rpdu\_3d\_pos\_y, and rpdu\_3d\_pos\_z. The number of bits used to represent atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 is Floor( Log2( gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 ) ). If atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 is not present, the number of bits used to represent atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 is Max( 0, gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 − gi\_geometry\_nominal\_2d\_bitdepth\_minus1 ) − 1.

[Ed. (All): Double check if the bit count for this information is correct.] [Ed. (PA): Bit count fixed]

**atgh\_num\_ref\_idx\_active\_override\_flag** equal to 1 specifies that the syntax element atgh\_num\_ref\_idx\_active\_minus1 is present for the current atlas tile group. atgh\_num\_ref\_idx\_active\_override\_flag equal to 0 specifies that the syntax element atgh\_num\_ref\_idx\_active\_minus1 is not present. If atgh\_num\_ref\_idx\_active\_override\_flag is not present, its value shall be inferred to be equal to 0.

**atgh\_num\_ref\_idx\_active\_minus1** is used for the derivation of the variable NumRefIdxActive as specified by Equation 7‑11 for the current atlas tile group. The value of atgh\_num\_ref\_idx\_active\_minus1 shall be in the range of 0 to 14, inclusive.

When the current atlas tile group is a P\_TILE\_GRP atlas tile group, atgh\_num\_ref\_idx\_active\_override\_flag is equal to 1, and atgh\_num\_ref\_idx\_active\_minus1 is not present, atgh\_num\_ref\_idx\_active\_minus1 is inferred to be equal to 0.

The variable NumRefIdxActive is derived as follows:

if( atgh\_type == P\_TILE\_GRP || atgh\_type == SKIP\_TILE\_GRP ) {  
 if( atgh\_num\_ref\_idx\_active\_override\_flag == 1 )  
 NumRefIdxActive = atgh\_num\_ref\_idx\_active\_minus1 + 1 (7‑11)  
 else {  
 if( num\_ref\_entries[ RlsIdx ] >= afps\_num\_ref\_idx\_default\_active\_minus1+ 1 )  
 NumRefIdxActive = afps\_num\_ref\_idx\_default\_active\_minus1+ 1  
 else  
 NumRefIdxActive = num\_ref\_entries[ RlsIdx ]  
 }  
 }   
 else   
 NumRefIdxActive = 0

The value of NumRefIdxActive − 1 specifies the maximum reference index for reference the atlas frame list that may be used to decode the current atlas tile group. When the value of NumRefIdxActive is equal to 0, no reference index for the reference atlas frame list may be used to decode the current atlas tile group.

#### Reference list structure semantics

**num\_ref\_entries**[ rlsIdx ] specifies the number of entries in the ref\_list\_struct( rlsIdx ) syntax structure. The value of num\_ref\_entries[ rlsIdx ] shall be in the range of 0 to asps\_max\_dec\_atlas\_frame\_buffering\_minus1+ 1, inclusive.

**st\_ref\_atlas\_frame\_flag**[ rlsIdx ][ i ] equal to 1 specifies that the i-th entry in the ref\_list\_struct( rlsIdx ) syntax structure is a short term reference atlas frame entry. st\_ref\_atlas\_frame\_flag[ rlsIdx ][ i ] equal to 0 specifies that the i-th entry in the ref\_list\_struct( rlsIdx ) syntax structure is a long term reference atlas frame entry. When not present, the value of st\_ref\_atlas\_frame\_flag[ rlsIdx ][ i ] is inferred to be equal to 1.

The variable NumLtrAtlasFrmEntries[ rlsIdx ] is derived as follows:

NumLtrAtlasFrmEntries[ rlsIdx ] = 0  
for( i = 0; i < num\_ref\_entries[ rlsIdx ]; i++ )  
 if( !st\_ref\_atlas\_frame\_flag[ rlsIdx ][ i ] ) (7‑12)  
 NumLtrAtlasFrmEntries[ rlsIdx ]++

**abs\_delta\_afoc\_st**[ rlsIdx ][ i ], when the i-th entry is the first short term reference atlas frame entry in ref\_list\_struct( rlsIdx ) syntax structure, specifies the absolute difference between the atlas frame order count values of the current atlas tile group and the atlas frame referred to by the i-th entry, or, when the i-th entry is a short term reference atlas frame entry but not the first short term reference atlas frame entry in the ref\_list\_struct( rlsIdx ) syntax structure, specifies the absolute difference between the atlas frame order count values of the atlas frames referred to by the i-th entry and by the previous short term reference atlas frame entry in the ref\_list\_struct( rlsIdx ) syntax structure.

The value of abs\_delta\_afoc\_st[ rlsIdx ][ i ] shall be in the range of 0 to 215 − 1, inclusive.

**strpf\_entry\_sign\_flag**[ rlsIdx ][ i ] equal to 1 specifies that i-th entry in the syntax structure ref\_list\_struct( rlsIdx ) has a value greater than or equal to 0. strpf\_entry\_sign\_flag[ rlsIdx ][ i ] equal to 0 specifies that the i-th entry in the syntax structure ref\_list\_struct( rlsIdx ) has a value less than 0. When not present, the value of strpf\_entry\_sign\_flag[ rlsIdx ][ i ] is inferred to be equal to 1.

The list DeltaAfocSt[ rlsIdx ][ i ] is derived as follows:

for( i = 0; i < num\_ref\_entries[ rlsIdx ]; i++ )   
 if( st\_ref\_atlas\_frame\_flag[ rlsIdx ][ i ] )   
 DeltaAfocSt[ rlsIdx ][ i ] =   
 (2 \* strpf\_entry\_sign\_flag[ rlsIdx ][ i ] − 1) \* abs\_delta\_afoc\_st[ rlsIdx ][ i ] (7‑13)  
 else  
 DeltaAfocSt[ rlsIdx ][ i ] = 0

**afoc\_lsb\_lt**[ rlsIdx ][ i ] specifies the value of the atlas frame order count modulo MaxAtlasFrmOrderCntLsb of the atlas frame referred to by the i-th entry in the ref\_list\_struct( rlsIdx ) syntax structure. The length of the afoc\_lsb\_lt[ rlsIdx ][ i ] syntax element is asps\_log2\_max\_atlas\_frame\_order\_cnt\_lsb\_minus4  + 4 bits.

### Atlas tile group data unit semantics

#### General atlas tile group data unit semantics

**atgdu\_patch\_mode**[ p ] indicates the patch mode for the patch with index p in the current atlas tile group. The permitted values for atgdu\_patch\_mode[ p ] are specified in Table 7‑6 for atlas tile groups with atgh\_type = I\_TILE\_GRP, Table 7‑7 for atlas tile groups with atgh\_type = P\_TILE\_GRP, and Table 7‑8 for atlas tile groups with atgh\_type = SKIP\_TILE\_GRP. A tile group with atgh\_type = SKIP\_TILE\_GRP implies that the entire tile group information is copied directly from the tile group with the same atgh\_address as that of the current tile group that corresponds to the first reference atlas frame in the RefAtlasFrmList.

Table 7‑6 Patch mode types for I\_TILE\_GRP type atlas tile groups

|  |  |  |
| --- | --- | --- |
| **atgdu\_patch\_mode** | **Identifier** | **Description** |
| 0 | I\_INTRA | Non-predicted Patch mode |
| 1 | I\_RAW | RAW Point Patch mode |
| 2 | I\_EOM | EOM Point Patch mode |
| 3-13 | I\_RESERVED | Reserved modes |
| 14 | I\_END | Patch termination mode |

Table 7‑7 Patch mode types for P\_TILE\_GRP type atlas tile groups

|  |  |  |
| --- | --- | --- |
| **atgdu\_patch\_mode** | **Identifier** | **Description** |
| 0 | P\_SKIP | Patch Skip mode |
| 1 | P\_MERGE | Patch Merge mode |
| 2 | P\_INTER | Inter predicted Patch mode |
| 3 | P\_INTRA | Non-predicted Patch mode |
| 4 | P\_RAW | RAW Point Patch mode |
| 5 | P\_EOM | EOM Point Patch mode |
| 6-13 | P\_RESERVED | Reserved modes |
| 14 | P\_END | Patch termination mode |

Table 7‑8 Patch mode types for SKIP\_TILE\_GRP type atlas tile groups

|  |  |  |
| --- | --- | --- |
| **atgdu\_patch\_mode** | **Identifier** | **Description** |
| 0 | P\_SKIP | Patch Skip mode |

#### Patch information data semantics

None

#### Patch data unit semantics

**pdu\_2d\_pos\_x**[ p ] specifies the x-coordinate of the top-left corner of the patch bounding box for patch p in the current atlas tile group, tileGroupIdx, expressed as a multiple of PatchPackingBlockSize.

**pdu\_2d\_pos\_y**[ p ] specifies the y-coordinate of the top-left corner of the patch bounding box for patch p in the current atlas tile group, tileGroupIdx, expressed as a multiple of PatchPackingBlockSize.

**pdu\_2d\_size\_x\_minus1**[ p ] plus 1 specifies the quantized width value of the patch with index p in the current atlas tile group, tileGroupIdx.

**pdu\_2d\_size\_y\_minus1**[ p ] plus 1 specifies the quantized height value of the patch with index p in the current atlas tile group, tileGroupIdx.

**pdu\_3d\_pos\_x**[ p ] specifies the shift to be applied to the reconstructed patch points in patch with index p of the current atlas tile group along the tangent axis. The value of pdu\_3d\_pos\_x[ p ] shall be in the range of 0 to Min( 2afps\_3d\_pos\_x\_bit\_count\_minus1 + 1, 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 ) − 1, inclusive. The number of bits used to represent pdu\_3d\_pos\_x[ p ] is afps\_3d\_pos\_x\_bit\_count\_minus1 + 1.

**pdu\_3d\_pos\_y**[ p ] specifies the shift to be applied to the reconstructed patch points in patch with index p of the current atlas tile group along the bitangent axis. The value of pdu\_3d\_pos\_y[ p ] shall be in the range of 0 to Min(2afps\_3d\_pos\_y\_bit\_count\_minus1 + 1, 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 ) − 1, inclusive. The number of bits used to represent pdu\_3d\_pos\_y[ p ] is afps\_3d\_pos\_y\_bit\_count\_minus1 + 1.

**pdu\_3d\_pos\_min\_z**[ p ] specifies the shift to be applied to the reconstructed patch points in patch with index p of the current atlas tile group along the normal axis, Pdu3dPosMinZ[ p ], as follows:

Pdu3dPosMinZ[ p ] = pdu\_3d\_pos\_min\_z[ p ] << atgh\_pos\_min\_z\_quantizer (7‑14)

The value of Pdu3dPosMinZ[ p ] shall be in the range of 0 to 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 – 1, inclusive.

The number of bits used to represent pdu\_3d\_pos\_min\_z[ p ] is equal to (gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 – atgh\_pos\_min\_z\_quantizer + 2).

**pdu\_3d\_pos\_delta\_max\_z**[ p ], if present, specifies the nominal maximum value of the shift expected to be present in the reconstructed bitdepth patch geometry samples, after conversion to their nominal representation, in patch with index p of the current atlas tile group along the normal axis, Pdu3dPosDeltaMaxZ[ p ], as follows:

Pdu3dPosDeltaMaxZ[ p ] = (pdu\_3d\_pos\_delta\_max\_z[ p ] == 0 ? 0:   
 (pdu\_3d\_pos\_delta\_max\_z[ p ] << atgh\_pos\_delta\_max\_z\_quantizer) – 1 (7‑15)

If pdu\_3d\_pos\_delta\_max\_z[ p ] is not present the value of Pdu3dPosDeltaMaxZ[ p ] is assumed to be equal to 2gi\_geometry\_nominal\_2d\_bitdepth\_minus1+1 – 1. When present, the value of Pdu3dPosDeltaMaxZ[ p ] shall be in the range of 0 to 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 − 1, inclusive.

The number of bits used to represent pdu\_3d\_pos\_delta\_max\_z[ p ] is equal to ( gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 – atgh\_pos\_delta\_max\_z\_quantizer + 2 ).

[Ed. (All): Include a note or create a section maybe in section 6 that would allow us to distinguish between local and global 3D coordinates and would clearly explain the relationship of these variables (as well as other 3D parameters) with the point cloud 3D space. Providing some figures could also be useful.]

**pdu\_projection\_id**[ p ] specifies the values of the projection mode and of the index of the normal to the projection plane for the patch with index p of the current atlas tile group. The value of pdu\_projection\_id[ p ] shall be in range of 0 to (asps\_extended\_projection\_enabled\_flag ? 17 : 5) , inclusive.

The variable PduProjectionPlane[ p ] is derived as follows:

PduProjectionPlane[ p ] =   
asps\_extended\_projection\_enabled\_flag ? (pdu\_projection\_id[ p ] >> 2) : pdu\_projection\_id[ p ]

Pdu45degreeProjectionRotationAxis[ p ] =   
 asps\_extended\_projection\_enabled\_flag? pdu\_projection\_id[ p ] & 0x03 : 0

The number of bits used to represent pdu\_projection\_id[ p ] is ( asps\_extended\_projection\_enabled\_flag ? 5 : 3)

**pdu\_orientation\_index**[ p ] indicates the index to Table 9‑1 of the patch orientation index for the patch with index p of the current atlas tile group. The number of bits used to represent pdu\_orientation\_index[ p ] is ( asps\_use\_eight\_orientations\_flag ? 3 : 1 )

**pdu\_lod\_enabled\_flag**[ p ] equal to 1 specifies that the LOD parameters are present for the current patch p. If pdu\_lod\_enabled\_flag[ p ] is equal to 0, no LOD parameters are present for the current patch. If pdu\_lod\_enabled\_flag[ p ] is not present, its value shall be inferred to be equal to 0.

**pdu\_lod\_scale\_x\_minus1**[ p ] specifies the LOD scaling factor to be applied to the local x coordinate of a point in a patch with index p of the current atlas tile group, prior to its addition to the patch coordinate Patch3dPosX[ p ]. If pdu\_lod\_scale\_x\_minus1[ p ] is not present, its value shall be inferred to be equal to 0.

**pdu\_lod\_scale\_y**[ p ] specifies the LOD scaling factor to be applied to the local y coordinate of a point in a patch with index p of the current atlas tile group, prior to its addition to the patch coordinate Patch3dPosY[ p ]. If pdu\_lod\_scale\_y[ p ] is not present, its value shall be inferred to be equal to 0.

#### Skip patch data unit semantics

None

#### Merge patch data unit semantics

**mpdu\_ref\_index**[ p ] specifies the atlas reference frame index, RefIdx , for the current patch. When mpdu\_ref\_index[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_override\_2d\_params\_flag**[ p ] specifies whether the 2d parameters for the current patch are present in the bitstream.

**mpdu\_2d\_pos\_x**[ p ] specifies the difference of the x-coordinate of the top-left corner of the patch bounding box of patch with index p in the current atlas tile group and of the x-coordinate of the top-left corner of the patch bounding box of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that is associated with the reference RefIdx, expressed as a multiple of PatchPackingBlockSize. When mpdu\_2d\_pos\_x[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_2d\_pos\_y**[ p ] specifies the difference of the y-coordinate of the top-left corner of the patch bounding box of patch with index p in the current atlas tile group and of the y-coordinate of the top-left corner of the patch bounding box of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that is associated with the reference RefIdx, expressed as a multiple of PatchPackingBlockSize. When mpdu\_2d\_pos\_x[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_2d\_delta\_size\_x**[ p ] specifies the difference of the width values of the patch with index p in the current atlas tile group and the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that is associated with the reference RefIdx. When mpdu\_2d\_delta\_size\_x[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_2d\_delta\_size\_y**[ p ] specifies the difference of the height values of the patch with index p in the current atlas tile group and the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that correponds to the reference RefIdx. When mpdu\_2d\_delta\_size\_y[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_override\_3d\_params\_flag**[ p ] specifies whether the 3d parameters for the current patch are present in the bitstream. When mpdu\_override\_3d\_params\_flag[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_3d\_pos\_x**[ p ] specifies the difference between the shift to be applied to the reconstructed patch points along the tangent axis of patch with index p in the current atlas tile group and of the shift to be applied to the reconstructed patch points along the tangent axis of patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to the reference RefIdx. The value of mpdu\_3d\_pos\_x[ p ] shall be in the range of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive. When mpdu\_3d\_pos\_x[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_3d\_pos\_y**[ p ] specifies the difference between the shift to be applied to the reconstructed patch points along the bitangent axis of patch with index p in the current atlas tile group and of the shift to be applied to the reconstructed patch points along the bitangent axis of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to RefIdx. The value of mpdu\_3d\_pos\_y[ p ] shall be in the range of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive. When mpdu\_3d\_pos\_y[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_3d\_pos\_min\_z**[ p ] specifies the difference between the shift to be applied to the reconstructed patch points along the normal axis of patch with index p in the current atlas tile group and of the shift to be applied to the reconstructed patch points along the normal axis of patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to RefIdx, Mpdu3dPosMinZ[ p ], as follows:

Mpdu3dPosMinZ[ p ] = mpdu\_3d\_pos\_min\_z[ p ] << atgh\_pos\_min\_z\_quantizer (7‑16)

The value of mpdu\_3d\_pos\_min\_z[ p ] shall be in the range of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive. When mpdu\_3d\_pos\_min\_z[ p ] is not present, it is inferred to be equal to 0.

**mpdu\_3d\_pos\_delta\_max\_z**[ p ], if present, specifies the difference between the nominal maximum value of the shift expected to be present in the reconstructed bitdepth patch geometry samples, after conversion to their nominal representation, in patch with index p of the current atlas tile group along the normal axis and of the nominal maximum value of the shift expected to be presented in the reconstructed bitdepth patch geometry samples, after conversion to their nominal representation of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to RefIdx, Mpdu3dPosDeltaMaxZ[ p ], as follows:

Mpdu3dPosDeltaMaxZ[ p ] = (mpdu\_3d\_pos\_delta\_max\_z[ p ] == 0 ? 0  
 : (mpdu\_3d\_pos\_delta\_max\_z[ p ] << atgh\_pos\_delta\_max\_z\_quantizer ) – 1 (7‑17)

If mpdu\_3d\_pos\_delta\_max\_z[ p ] is not present the value of Mpdu3dPosDeltaMaxZ[ p ] is assumed to be equal to 2gi\_geometry\_nominal\_2d\_bitdepth\_minus1+1 – 1.When present, the value of mpdu\_3d\_pos\_delta\_max\_z[ p ] shall be in the range of of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive.

**mpdu\_override\_plr\_flag**[ patchIdx ] specifies whether the point local reconstruction parameters may be overwritten by new parameters that are present in the bistream for the patch.

#### Inter patch data unit semantics

**ipdu\_ref\_index**[ p ] specifies the atlas reference frame index, RefIdx , for the current patch. When ipdu\_ref\_index[ p ] is not present, it is inferred to be equal to 0.

**ipdu\_patch\_index**[ p ] specifies the index, PredIdx, of the patch in the atlas tile group with the same address as the current tile group adress in the atlas frame that corresponds to index RefIdx in the current reference atlas frame list.

[Ed. (ATA): Semantic description above need to be re-written. One alternative could be: “when p is equal to 0, specifies, PredIdx, of the patch in the atlas tile group with the same address as the current tile group address in the atlas frame that corresponds to index RefIdx  in the current reference atlas frame list. When p is larger than 0, ipdu\_patch\_index[ p ] specifies the difference of the index values of the patch with index p and the reference patch with index (PredIdx + 1)”. After discussing with AMT, even in case of p = 0, this is a difference as well hence the definition can be unified. There are some discussions items still e.g. for case p = 0 ue(v) vs. se(v) for p > 0 .]

**ipdu\_2d\_pos\_x**[ p ] specifies the difference of the x-coordinate of the top-left corner of the patch bounding box of patch with index p in the current atlas tile group and of the x-coordinate of the top-left corner of the patch bounding box of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that is associated with the reference RefIdx, expressed as a multiple of PatchPackingBlockSize.

**ipdu\_2d\_pos\_y**[ p ] specifies the difference of the y-coordinate of the top-left corner of the patch bounding box of patch with index p in the current atlas tile group and of the y-coordinate of the top-left corner of the patch bounding box of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that is associated with the reference RefIdx, expressed as a multiple of PatchPackingBlockSize.

**ipdu\_2d\_delta\_size\_x**[ p ] specifies the difference of the width values of the patch with index p in the current atlas tile group and the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that is associated with the reference RefIdx.

**ipdu\_2d\_delta\_size\_y**[ p ] specifies the difference of the height values of the patch with index p in the current atlas tile group and the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to the reference RefIdx.

**ipdu\_3d\_pos\_x**[ p ] specifies the difference between the shift to be applied to the reconstructed patch points along the tangent axis of patch with index p in the current atlas tile group and of the shift to be applied to the reconstructed patch points along the tangent axis of patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to the reference RefIdx. The value of ipdu\_3d\_pos\_x[ p ] shall be in the range of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive.

**ipdu\_3d\_pos\_y**[ p ] specifies the difference between the shift to be applied to the reconstructed patch points along the bitangent axis of patch with index p in the current atlas tile group and of the shift to be applied to the reconstructed patch points along the bitangent axis of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to RefIdx. The value of ipdu\_3d\_pos\_y[ p ] shall be in the range of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive.

**ipdu\_3d\_pos\_min\_z**[ p ] specifies the difference between the shift to be applied to the reconstructed patch points along the normal axis of patch with index p in the current atlas tile group and of the shift to be applied to the reconstructed patch points along the normal axis of patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to RefIdx, Ipdu3dPosMinZ[ p ], as follows:

Ipdu3dPosMinZ[ p ] = ipdu\_3d\_pos\_min\_z[ p ] << atgh\_pos\_min\_z\_quantizer (7‑18)

The value of ipdu\_3d\_pos\_min\_z[ p ] shall be in the range of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive.

**ipdu\_3d\_pos\_delta\_max\_z**[ p ], if present, specifies the difference between the nominal maximum value of the shift expected to be present in the reconstructed bitdepth patch geometry samples, after conversion to their nominal representation, in patch with index p of the current atlas tile group along the normal axis and of the nominal maximum value of the shift expected to be presented in the reconstructed bitdepth patch geometry samples, after conversion to their nominal representation of the patch with index PredIdx in the atlas tile group with the same address as the current tile group in the atlas frame that corresponds to RefIdx, Ipdu3dPosDeltaMaxZ[ p ], as follows:

Ipdu3dPosDeltaMaxZ[ p ] = (ipdu\_3d\_pos\_delta\_max\_z[ p ] == 0 ? 0  
 : ( ipdu\_3d\_pos\_delta\_max\_z[ p ] << atgh\_pos\_delta\_max\_z\_quantizer ) – 1 (7‑19)

If ipdu\_3d\_pos\_delta\_max\_z[ p ] is not present the value of Ipdu3dPosDeltaMaxZ[ p ] is assumed to be equal to 2gi\_geometry\_nominal\_2d\_bitdepth\_minus1+1 – 1.When present, the value of ipdu\_3d\_pos\_delta\_max\_z[ p ] shall be in the range of of (−2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 + 1) to (2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1+1 − 1), inclusive.

[Ed. (AMT): When not present, actually what we need to set to the default is the actual value, not the diff. May need some cleanup, but could be fixed later]

[Ed. (RJ): Please refer to comment regarding ipdu\_projection\_mode in clause 7.3.32: Inter Patch data unit syntax]

[Ed. (AMT): What should the value and definition be for multiple maps of value > 2? How about also for the case when the data are interleaved? This seems to be currently "broken"]

#### RAW patch data unit semantics

**rpdu\_patch\_in\_auxiliary\_video\_flag**[ patchIdx ] specifies whether the geometry and attribute data associated with the RAW coded patch with index patchIdx in the current atlas tile group are encoded in an auxiliary video compared to those of the intra and inter coded patches. If rpdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is equal to 0, the geometry and attribute data associated with the RAW coded patch patchIdx in the current atlas tile group are encoded in the same video as those of the intra and inter coded patches. If rpdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is equal to 1, the geometry and attribute data associated with the RAW coded patch patchIdx in the current atlas tile group are encoded in an auxiliary video from those of the intra and inter coded patches. If rpdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is not present, its value shall be inferred to be equal to 0.

**rpdu\_2d\_pos\_x**[ patchIdx ] specifies the x-coordinate of the top-left corner of the patch bounding box size for RAW coded patch patchIdx in the current atlas tile group, tileGroupIdx, expressed as a multiple of PatchPackingBlockSize.

**rpdu\_2d\_pos\_y**[ patchIdx ] specifies the y-coordinate of the top-left corner of the patch bounding box size for RAW coded patch patchIdx in the current atlas tile group, tileGroupIdx, expressed as a multiple of PatchPackingBlockSize.

**rpdu\_2d\_size\_x\_minus1**[ patchIdx ] plus 1 specifies the width value of the RAW coded patch with index patchIdx in the current atlas tile group.

**rpdu\_2d\_size\_y\_minus1**[ patchIdx ] plus 1 specifies the height value of the RAW coded patch with index patchIdx in the current atlas tile group.

**rpdu\_3d\_pos\_x**[ patchIdx ] specifies the shift to be applied to the reconstructed RAW patch points in the patch with index patchIdx of the current atlas tile group along the tangent axis. The number of bits used to represent rpdu\_3d\_pos\_x[ patchIdx ] is ( atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 + 1 ).

**rpdu\_3d\_pos\_y**[ patchIdx ] specifies the shift to be applied to the reconstructed RAW patch points in the patch with index patchIdx of the current atlas tile group along the bitangent axis. The number of bits used to represent rpdu\_3d\_pos\_y[ patchIdx ] is ( atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 + 1 ).

**rpdu\_3d\_pos\_z**[ patchIdx ] specifies the shift to be applied to the reconstructed RAW patch points in the patch with index patchIdx of the current atlas tile group along the normal axis. The number of bits used to represent rpdu\_3d\_pos\_z[ patchIdx ] is ( atgh\_raw\_3d\_pos\_axis\_bit\_count\_minus1 + 1 ).

**rpdu\_points\_minus1**[ patchIdx ] plus 1 specifies the number of RAW points present in the RAW coded patch with index patchIdx  in the current atlas tile group. The value of rpdu\_points\_minus1[ patchIdx ] shall be in the range of 0 to (rpdu\_2d\_size\_x\_minus1[ patchIdx ] + 1) \* (rpdu\_2d\_size\_y\_minus1 [ patchIdx ] + 1) − 1, inclusive.

[Ed. (PA) (RJ): The last sentence does not make sense. Is it intended to specify a default value of 0?]

[Ed. (AMT): RAW points assume 3 elements. Should this be generalized for attributes with more components?]

[Ed. (JL): Removed the sentence about the range of possible valuesfor RAW points because it was incorrect. The size of the RAW patch bounding box depends on the number of points and the “interleaving” approach & chroma sampling.][Ed. (AMT): This should be fixed][Ed. (PA) We may add a limit to rpdu\_points according to size of RAW patch]

#### EOM patch data unit semantics

**epdu\_patch\_in\_auxiliary\_video\_flag**[ patchIdx ] specifies whether the attribute data associated with the current EOM patch with index patchIdx in the current atlas tile group are encoded in an auxiliary video compared to those of the intra and inter coded patches. If epdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is equal to 0, the attribute data associated with the EOM patch with index patchIdx in the current atlas tile group are encoded in the same video as those of the intra and inter coded patches. If epdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is equal to 1, the attribute data associated with the current EOM patch with index patchIdx in the current atlas tile group are encoded in an auxiliary video from those of the intra and inter coded patches. If epdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is not present, its value shall be inferred to be equal to 0.

**epdu\_2d\_pos\_x**[ patchIdx ] specifies the x-coordinate of the top-left corner of the patch bounding box size for the current EOM patch patchIdx in the current atlas tile group, tileGroupIdx, expressed as a multiple of PatchPackingBlockSize. If the value of epdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is equal to 0, the value of epdu\_2d\_pos\_x[ patchIdx ] shall be in the range of 0 to ( TileGroupWidth[ tileGroupIdx ] / PatchPackingBlockSize ) − 1, inclusive.

**epdu\_2d\_pos\_y**[ patchIdx ] specifies the y-coordinate of the top-left corner of the patch bounding box size for the current EOM patch with index patchIdx in the current atlas tile group, tileGroupIdx, expressed as a multiple of PatchPackingBlockSize. If the value of epdu\_patch\_in\_auxiliary\_video\_flag[ patchIdx ] is equal to 0, the value of epdu\_2d\_pos\_y[ patchIdx ] shall be in the range of 0 to ( TileGroupHeight[ tileGroupIdx ] / PatchPackingBlockSize ) − 1.

**epdu\_2d\_size\_x\_minus1**[ patchIdx ] plus 1 specifies the width value of the current EOM patch with index patchIdx in the current atlas tile group.

**epdu\_2d\_size\_y\_minus1**[ patchIdx ] plus 1 specifies the height value of the current EOM patch with index patchIdx in the current atlas tile group.

**epdu\_patch\_count\_minus1**[ patchIdx ] plus 1 indicates the number of patches that contain EOM points in the occupancy map for the current EOM patch with index patchIdx in the current atlas tile group. The value of epdu\_patch\_count\_minus1[ patchIdx ] shall be in the range of 0 to 2afps\_eom\_number\_of\_patch\_bit\_count\_minus1 + 1 – 1, inclusive. The number of bits used to represent epdu\_patch\_count\_minus1[ patchIdx ] is afps\_eom\_number\_of\_patch\_bit\_count\_minus1 + 1.

[Ed. (PA): TMC2 indicates a fixed u(8) while spec indicates an undefined u(v).]

**epdu\_associated\_patch\_idx**[ patchIdx ][ i ] indicates the corresponding patch index in raster scan order, of a patch that contain EOM points in the occupancy map, of the i-th patch specified in the current EOM patch with index patchIdx in the current atlas tile group and which is associated with a number of EOM attribute points as indicated by the syntax element epdu\_points[ patchIdx ][ i ]. The value of epdu\_associated\_patch\_idx[ patchIdx ][ i ] shall be in the range 0 to 2afps\_eom\_number\_of\_patch\_bit\_count\_minus1 + 1 – 1, inclusive.

[Ed. (PA): TMC2 indicates a fixed u(8) while spec indicates an ue(v).]

**epdu\_points**[ patchIdx ][ i ] specifies the number of EOM points associated with the j-th patch specified in the current EOM patch with index patchIdx in the current atlas tile group that contain EOM attributes. The value of epdu\_points[ patchIdx ][ i ] shall be in the range of 0 to 2afps\_eom\_max\_bit\_count\_minus1 + 1 – 1, inclusive. The number of bits used to represent epdu\_points[ patchIdx ][ i ] is afps\_eom\_max\_bit\_count\_minus1 + 1. When not present epdu\_points[ patchIdx ][ i ] shall be inferred to be equal to 0.

[Ed. (PA): TMC2 indicates an ue(v) while spec indicates an u(v).]

[Ed. (PA):May we want to signal epdu\_point equal to 0?]

#### Point local reconstruction data semantics

The variable BlockCount is derived as follows:

BlockCount= Patch2dSizeX[ patchIdx ] \* Patch2dSizeY[ patchIdx ]

with the variables Patch2dSizeX[ patchIdx ] and Patch2dSizeY[ patchIdx ] derived as specified in clause 8.4.4.1.

**plrd\_level**[ mapIdx ][ patchIdx ] equal to 0 indicates that point local reconstruction data information is present in the bitstream for each block of the map mapIndex of the patch patchIndex in the current atlas tile group. plrd\_level[ mapIdx ][ patchIdx ] equal to 1 indicates that point local reconstruction data information is present in the bitstream for the entire map mapIndex of the patch patchIdx in the current atlas tile group. In this case, all blocks of the map mapIdx shall use the same point local reconstruction data. When plrd\_level[ mapIdx ][ patchIdx ] is equal to 0, the syntax elements plrd\_present\_block\_flag[ mapIdx ][ patchIdx ][ i ] and plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] are present in the bitstream. When plrd\_level[ mapIdx ][ patchIdx ] is equal to 1, the syntax elements plrd\_present\_flag[ mapIdx ][ patchIdx ] and plrd\_mode\_minus1[ mapIdx ][ patchIdx ].

**plrd\_present\_block\_flag**[ mapIdx ][ patchIdx ][ i ] equal to 1 indicates that the syntax element plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] is present in the bitstream for block i of the map mapIdx of the patch patchIdx in the current atlas tile group. plrd\_present\_block\_flag[ mapIdx ][ patchIdx ][ i ] equal to 0 indicates that the syntax element plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] is not present in the bitstream for block i of the map mapIdx of the patch patchIdx in the current atlas tile group.

**plrd\_block\_mode\_minus1**[ mapIdx ][ patchIdx ][ i ] plus 1 indicates the point local reconstruction mode selected in the list of point local reconstruction modes of size plri\_number\_of\_modes\_minus1[ mapIdx ] plus 1 and defined by the syntax elements plri\_interpolate\_flag[ mapIdx ][ plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ], plri\_filling\_flag[ mapIdx ][plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ], plri\_minimum\_depth[ mapIdx ][ plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ] and plri\_neighbour\_minus1[ mapIdx ][ plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ] for block i of the map mapIdx of the patch patchIdx in the current atlas tile group. plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] shall be in the range of 0 to plri\_number\_of\_modes\_minus1[ mapIdx ], inclusive.

NOTE – When PLR mode is equal to 0 at the block or patch level, it is not signalled in the bitstream. The value 0 corresponds to a default mode value of PLR which is implemented by the values of the following syntax elements plri\_interpolate\_flag[ mapIdx ][ 0 ], plri\_filling\_flag[ mapIdx ][ 0 ], plri\_minimum\_depth[ mapIdx ][ 0 ] and plri\_neighbout\_minus1[ mapIdx ][ 0 ].

[Ed. (ATA/Danillo): should this be index to the mode list that has been created in the asps. Based on the above definition mode 0 may be exclude?]

[Ed. (PA): The intent is to signal only when mode is different from mode 0. Mode 0 is a default/initialization mode. This is aligned with the S/W. Corrections brought in this version -- see clause 9.4.2. Added a note to clarify. ]

**plrd\_present\_flag**[ mapIdx ][ patchIdx ] equal to 1 indicates that the syntax element plrd\_mode\_minus1[ mapIdx ][ patchIdx ] is present in the bitstream for the map mapIdx of the patch patchIdx in the current atlas tile group. plrd\_present\_flag[ mapIdx ][ patchIdx  ] equal to 0 indicates that the syntax element plrd\_block\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] is not present in the bitstream for the map mapIdx of the patch patchIdx in the current atlas tile group.

**plrd\_mode\_minus1**[ mapIdx ][ patchIdx ] plus 1 indicates the point local reconstruction mode selected in the list of point local reconstruction modes of size plri\_number\_of\_modes\_minus1[ mapIdx ] plus 1 and defined by the syntax elements plri\_interpolate\_flag[ mapIdx ][ plrd\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ], plri\_filling\_flag[ mapIdx ][plrd\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ], plri\_minimum\_depth[ mapIdx ][ plrd\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ] and plri\_neighbour\_minus1[ mapIdx ][ plrd\_mode\_minus1[ mapIdx ][ patchIdx ][ i ] + 1 ] for all blocks of the map mapIdx of the patch patchIdx in the current atlas tile group. plrd\_mode\_minus1[ mapIdx ][ patchIdx ] shall be in the range 0 to plri\_number\_of\_modes\_minus1[ mapIdx ], inclusive.

### Supplemental enhancement information message semantics

Each SEI message consists of the variables specifying the type payloadType and size payloadSize of the SEI message payload. SEI message payloads are specified in Annex E. The derived SEI message payload size payloadSize is specified in bytes and shall be equal to the number of bytes in the SEI message payload.

**sm\_payload\_type\_byte** is a byte of the payload type of an SEI message.

**sm\_payload\_size\_byte** is a byte of the payload size of an SEI message.

# Decoding process

[Ed. (All): Add a section before the decoding section that includes an overview of the complete point cloud decoding & reconstruction. The name decoding may be confusing.]

[Ed. (VZ): As a general comment I would suggest reshaping the order of decoding description: Aux, OM, Geo, and Attr.]

## General decoding process

Input to this process is a V-PCC bitstream or a collection of V-PCC sub-bitstream components.

Output of this process is a set of decoded video streams, corresponding to the occupancy, geometry, and attribute, if available, information of the point cloud sequence, as well as the decoded atlas information, which are needed to perform the 3D reconstruction process specified in clause 9. The decoding process may include additional processing steps required to convert the decoded video streams to a nominal format, i.e. a nominal resolution, bitdepth, chroma representation, and frame rate. Such conversion process is outside of the scope of this specification, however for purposes of conformance to point B a process is described in Section 9.2.

The decoding process may already include the 3D reconstruction process, as specified in clause 9. For such a case, the output of the decoding process is a sequence of decoded point cloud frames. A point cloud frame consists of a set of points with (x, y, z) coordinates. Optionally, it may also contain one or more set of attributes, as shown in Table 7‑2, that shall be associated with each point.

[Ed (All): The above statement should be moved to the profiles/levels Annex]

The decoding process is specified such that all decoders that conform to a specific image/video codec and profile, tier, and level will produce numerically identical outputs 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 outputs to those produced by the process described herein (with the correct output order or output timing, as specified) conforms to the decoding process requirements of this document.

NOTE – For the purpose of best-effort decoding, a decoder that conforms to a particular profile at a given tier and level may additionally decode some bitstreams conforming to a different tier, level, or profile without necessarily using a decoding process that produces numerically identical decoded outputs to those produced by the process specified herein (without claiming conformance to the other profile, tier, and level).

[Ed. (AMT): Do we need to mention anything about best-effort decoding? This is coming from the HEVC specification where this support was requested. For information, I have made this into a note, as is also the case in the HEVC spec, since this is not a necessary process.]

The decoding process operates as follows:

1. If the input is a V-PCC bitstream then the V-PCC bitstream is parsed and demultiplexed into several sub-bitstreams, each one corresponding to a different set of information. In particular, the V-PCC bitstream is demultiplexed into occupancy map, geometry, and attributes, if present, video sub-bitstreams, and in an atlas sub-bitstream.
2. The demultiplexed sub-bitstreams in the previous process or the sub-bitstream if the input was a collection of V-PCC sub-bitstream components are then decoded as specified in the subsequent steps.
3. The occupancy map video decoding process as specified in clause 8.5 is invoked with the sub-bitstream corresponding to the occupancy map information as the input and the decoded occupancy map video frames, as the output. Each decoded occupancy map video frame is represented as oFrame[ y ][ x ], where, y and x are the row and column indices, respectively. The variable y is in the range of 0 to asps\_frame\_height – 1, inclusive and the variable x is in the range of 0 to asps\_frame\_width – 1, inclusive.
4. The geometry video decoding process, as specified in clause 8.2, is invoked.
5. For every present attribute video sub-bitstreams in the bitstream, the attribute video decoding process as specified in clause 8.3, is invoked .
6. The atlas data decoding process, as specified in clause 8.4, is invoked with the sub-bitstream corresponding to the encoded atlas data information as the input and the decoded atlas data information as the output.

[Ed. (AMT): Atlas data should be changed to refer to NAL units here and in other parts.]

[Ed. (SS): Decoding steps do not necessarily have to follow order 1-4 but could be rearranged or performed in parallel, depending on bit stream configuration and access][Ed. (All): Add note.]

[Ed. (AMT): Style for the different steps seems different. in the Occupancy and atlas sequence case we specify inputs and outputs, but we do not do so for the geometry and attributes. For the Occupancy we also specify the name of the decoded output. Should we harmonize the steps for consistency? It might be preferred actually that some of the details are mentioned in the subclauses than here.]

[Ed. (AMT): Note that for the occupancy, oFrame is specified based on the nominal resolution, not the video stream resolution. That seems to also go into "reconstruction" and not only decoding since there are multiple ways to interpolate the occupancy (if subsampling is used), while if the occupancy is not using binary encoding, the "thresholding" process could also be non-normative in that scenario. Again, likely this could be handled better if the text above was moved into section 8.5.]

[Ed. (AMT): Should section 8.5 about occupancy decoding be moved also ahead of the others?]

[Ed. (ATA): Is it better to use atlas NAL unit instead of atlas data?]

## Occupancy video decoding process

The occupancy video decoding process for the current atlas, with index atlasIdx, is invoked using the occupancy video sub-bitstream present in the V-PCC bitstream as the input. The occupancy video sub-bitstream is associated with a codec as specified by oi\_occupancy\_codec\_id[ atlasIdx ]. Outputs of this process are the decoded and output ordered occupancy video frames, decOccFrame[ orderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, decOccBitdepth[ orderIdx ], luma width, decOccWidth[ orderIdx ], and luma height, decOccHeight[ orderIdx ], where orderIdx is the output order index of the decoded occupancy frames, compIdx corresponds to the colour component index and is equal to 0, y is the row index in the decoded frame and is in the range of 0 to decOccHeight[ orderIdx ] − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to decOccWidth[ orderIdx ] − 1, inclusive.

## Geometry video decoding process

The geometry video decoding process for the current atlas, with index atlasIdx, is invoked using each geometry video sub-bitstreams present in the V-PCC bitstream as the input. Each geometry video sub-bitstream is associated with a separate value of vuh\_map\_index and an associated codec specified by gi\_geometry\_codec\_id[ atlasIdx ]. Outputs of this process are the decoded and output ordered geometry video frames, decGeoFrame[ mapIdx ][ orderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, decGeoBitdepth[ mapIdx ][ orderIdx ], luma width, decGeoWidth[ mapIdx ][ orderIdx ], and luma height, decGeoHeight[ mapIdx ][ orderIdx ], where mapIdx corresponds to the map index of each frame and is in the range of 0 to vps\_multiple\_map\_streams\_present\_flag[ atlasIdx ] ? vps\_map\_count\_minus1[ atlasIdx ] : 0, inclusive, orderIdx is the output order index of the decoded geometry frames, compIdx corresponds to the colour component index and is equal to 0, y is the row index in the decoded frame and is in the range of 0 to decGeoHeight[ mapIdx ][ orderIdx ] − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to decGeoWidth[ mapIdx ][ orderIdx ] − 1, inclusive.

## Attribute video decoding process

The attribute video decoding process for the current atlas atlasIdx is performed as follows:

– If ai\_attribute\_count[ atlasIdx ] is equal to 0, no attribute video frames are decoded and no attribute information is associated with the final, reconstructed point cloud.

– Otherwise (if ai\_attribute\_count[ atlasIdx ]  is not equal to 0), for each attribute with index attrIdx, where attrIdx is in the range of 0 to ai\_attribute\_count[ atlasIdx ] − 1, inclusive, (ai\_attribute\_dimension\_partitions\_minus1[ atlasIdX ][ attrIdx ] + 1) video decoding processes are invoked, each with a value of vuh\_attribute\_index equal to attrIdx and a different value of vuh\_attribute\_partition\_index associated with it and the associated codec specified by ai\_attribute\_codec\_id[ atlasIdx ][ attrIdx ] as the input. Outputs of this process are the decoded and output ordered attribute video frames, decAttrFrame[ attrIdx ][ mapIdx ][ partIdx ][ orderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, decAttrBitdepth[ attrIdx ][ mapIdx ][ partIdx ][ orderIdx ], width, decAttrWidth[ attrIdx ][ mapIdx ][ partIdx ][ orderIdx ], and height, decAttrHeight[ attrIdx ][ mapIdx ][ partIdx ][ orderIdx ], information, where partIdx corresponds to the attribute partition index and is in the range of 0 to ai\_attribute\_dimension\_partitions\_minus1[ atlasIdx ][ attrIdx ], inclusive, mapIdx corresponds corresponds to the map index of each frame and is in the range of 0 to vps\_multiple\_map\_streams\_present\_flag[ atlasIdx ] ? vps\_map\_count\_minus1[ atlasIdx ] : 0, inclusive, orderIdx is the output order index of the decoded attribute frames, compIdx corresponds to the attribute partition component index and is in the range of 0 to ai\_attribute\_partition\_channels\_minus1[ atlasIdx ][ attrIdx ][ partIdx ], inclusive, y is in the range of 0 to decAttrHeight[ attrIdx ][ mapIdx ][ partIdx ][ orderIdx ] − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to decAttrWidth[ attrIdx ][ mapIdx ][ partIdx ][ orderIdx ] − 1, inclusive.

## Atlas data decoding process

### General atlas data decoding process

Input to this process is the bitstream corresponding to the atlas data unit of the V-PCC bitstream.

Outputs of this process are the decoded atlas frames, as well any associated geometry and/or attribute parameters contained in the atlas data unit, ordered in increasing order according to their corresponding atlas frame order count number.

The decoding process operates as follows for the current atlas tile group, CurrPatchFrame, being decoded:

1. The decoding of atlas data units is specified in clause 8.4.2.
2. The processes in clause 8.4.3 specify the following decoding processes using syntax elements in the atlas frame header layer and above:

– Variables and functions relating to atlas frame order count are derived as specified in clause 8.4.3.1.

– At the beginning of the decoding process for each atlas frame, the reference atlas frame list construction process specified in clause 8.4.3.2 is invoked for the derivation of the reference atlas frame list, RefAtlasFrmList.

– The reference atlas frame marking process in clause 8.4.3.3 is invoked, wherein reference atlas frames may be marked as "unused for reference" or "used for long-term reference".

1. The processes in clause 8.4.4 specify the patch decoding processes according to the patch mode as follows:

– Decoding of intra coded patches is specified in clause 8.4.4.2.

– Decoding of skip mode coded patches is specified in clause 8.4.4.3

– Decoding of merge mode coded patches is specified in clause 8.4.4.4.

– Decoding of inter coded patches is specified in clause 8.4.4.5.

– Decoding of RAW coded patches is specified in clause 8.4.4.6.

– Decoding of EOM coded patches is specified in clause 8.4.4.7.

1. The process in clause 8.4.5 specifies the creation of the block to patch map after the decoding of all patches within the current atlas tile group.
2. After the current atlas tile group have been decoded, it is marked as "used for short-term reference".

### Atlas data unit decoding process

Inputs to this process are atlas data units of the current atlas tile group and their associated parameter set units.

Outputs of this process are the parsed syntax structures encapsulated within these atlas data units.

The decoding process for each atlas data unit extracts the syntax structure from the atlas data unit and then parses the syntax structure.

### Atlas data header decoding process

#### Atlas frame order count derivation process

Output of this process is AtlasFrmOrderCntVal, the atlas frame order count of the current atlas tile group, with index AtlasFrmIndex.

Atlas frame order counts are used to identify output order of atlas frames, as well as for decoder conformance checking.

Each coded atlas frame is associated with an atlas frame order count variable, denoted as AtlasFrmOrderCntVal.

When there is no available reference atlas frame in the reference atlas frame buffer, the variables prevAtlasFrmOrderCntLsb and prevAtlasFrmOrderCntMsb are derived as follows:

* Let prevAtlasFrm be the previous atlas frame in decoding order.
* The variable prevAtlasFrmOrderCntLsb is set equal to the atlas frame order count LSB value of prevAtlasFrm.
* The variable prevAtlasFrmOrderCntMsb is set equal to AtlasFrmOrderCntMsb of prevAtlasFrm.

The variable AtlasFrmOrderCntMsb of the current atlas tile group is derived as follows:

* If there is no available reference atlas frame in the reference atlas frame buffer, AtlasFrmOrderCntMsb is set equal to 0.
* Otherwise, AtlasFrmOrderCntMsb is derived as follows:

if( ( atgh\_atlas\_frm\_order\_cnt\_lsb <  prevAtlasFrmOrderCntLsb ) &&  
 ( ( prevAtlasFrmOrderCntLsb − atgh\_atlas\_frm\_order\_cnt\_lsb) >=   
 ( MaxAtlasFrmOrderCntLsb / 2 ) ) )

AtlasFrmOrderCntMsb = prevAtlasFrmOrderCntMsb + MaxAtlasFrmOrderCntLsb (8‑1)

else if( ( atgh\_atlas\_frm\_order\_cnt\_lsb > prevAtlasFrmOrderCntLsb ) &&  
 ( ( atgh\_atlas\_frm\_order\_cnt\_lsb − prevAtlasFrmOrderCntLsb ) >  
 ( MaxAtlasFrmOrderCntLsb / 2 ) ) )

AtlasFrmOrderCntMsb = prevAtlasFrmOrderCntMsb – MaxAtlasFrmOrderCntLsb

else

AtlasFrmOrderCntMsb = prevAtlasFrmOrderCntMsb

AtlasFrmOrderCntVal is derived as follows:

AtlasFrmOrderCntVal = AtlasFrmOrderCntMsb + atgh\_atlas\_frm\_order\_cnt\_lsb (8‑2)

The value of AtlasFrmOrderCntVal shall be in the range of −231 to 231 − 1, inclusive. In one CAS, the AtlasFrmOrderCntVal values for any two coded atlas frames shall not be the same.

At any moment during the decoding process, the values of AtlasFrmOrderCntVal & ( MaxLtAtlasFrmOrderCntLsb − 1 ) for any two reference atlas frames in the DPB shall not be the same.

The function AtlasFrmOrderCnt( pFrmX) is specified as follows:

AtlasFrmOrderCnt( pFrmX) = AtlasFrmOrderCntVal of the atlas frame pFrmX (8‑3)

The function DiffAtlasFrmOrderCnt( pFrmA, pFrmB) is specified as follows:

DiffAtlasFrmOrderCnt( pFrmA, pFrmB) =  
 AtlasFrmOrderCnt( pFrmA) – AtlasFrmOrderCnt( pFrmB) (8‑4)

The bitstream shall not contain data that result in values of DiffAtlasFrmOrderCnt( pFrmA, pFrmB) used in the decoding process that are not in the range of −215 to 215 − 1, inclusive.

NOTE – Let X be the current atlas tile group and Y and Z be two other atlas frames in the same CAS, Y and Z are considered to be in the same output order direction from X when both DiffAtlasFrmOrderCnt( X, Y ) and DiffAtlasFrmOrderCnt( X, Z ) are positive or both are negative.

#### Reference atlas frame list construction process

This process is invoked at the beginning of the decoding process for each atlas frame.

Reference atlas frames are addressed through reference indices. A reference index is an index into a reference atlas frame list. When decoding an I\_TILE\_GRP atlas tile group, no reference atlas frame list is used in decoding of the atlas tile group data. When decoding a P\_TILE\_GRP atlas tile group, a single reference atlas frame list, RefAtlasFrmList, is used in decoding of the atlas frame data.

At the beginning of the decoding process for each atlas frame, the reference atlas frame list RefAtlasFrmList is derived. The reference atlas frame list is used in marking reference atlas frames as specified in clause 8.4.3.3 or in decoding of the atlas tile group data.

The reference atlas frame list RefAtlasFrmList is constructed as follows:

for( j = 0, afocBase = AtlasFrmOrderCntVal; j < num\_ref\_entries[ RlsIdx ]; j++) { (8‑5)  
 if( st\_ref\_atlas\_frame\_flag[ RlsIdx ][ j ] ) {  
 RefAtlasFrmAfocList[ j ] = afocBase − DeltaAfocSt [ RlsIdx ][ j ]  
 if( reference pfA exists in the DAFB with AtlasFrmOrderCntVal equal to RefAtlasFrmAfocList[ j ] )  
 RefAtlasFrmList[ j ] = pfA  
 else  
 RefAtlasFrmList[ j ] = "no reference atlas frame"  
 afocBase = RefAtlasFrmAfocList[ j ]  
 } else {  
 if(reference pfA exists in the DAFB with AtlasFrmOrderCntVal & ( MaxLtAtlasFrmOrderCntLsb − 1 )  
 equal to FullAtlasFrmOrderCntLsbLt[ RlsIdx ][ j ] )  
 RefAtlasFrmList[ j ] = pfA  
 else  
 RefAtlasFrmList[ j ] = "no reference atlas frame"  
 }  
 }

The first NumRefIdxActive[ i ] entries in RefAtlasFrmList are referred to as the active entries in RefAtlasFrmList and the other entries in RefAtlasFrmList are referred to as the inactive entries in RefAtlasFrmList.

It is a requirement of bitstream conformance that the following constraints apply:

* num\_ref\_entries[ RlsIdx ] shall not be less than NumRefIdxActive[ i ].
* The atlas frame referred to by each active entry in RefAtlasFrmList shall be present in the DAFB.
* The atlas frame referred to by each entry in RefAtlasFrmList shall not be the current atlas frame.
* A short term reference atlas frame entry and a long term reference atlas frame entry in RefAtlasFrmList of an atlas frame shall not refer to the same atlas frame.
* There shall be no long term reference atlas frame entry in RefAtlasFrmList for which the difference between the AtlasFrmOrderCntVal of the current atlas tile group and the AtlasFrmOrderCntVal of the atlas frame referred to by the entry is greater than or equal to 224.
* Let setOfRefAtlasFrms be the set of unique atlas frames referred to by all entries in RefAtlasFrmList. The number of atlas frames in setOfRefAtlasFrms shall be less than or equal to asps\_max\_dec\_atlas\_frame\_buffering\_minus1.
* The atlas frames referred to by each active entry in RefAtlasFrmList shall have exactly the same tile group partitioning as the current atlas frame.
* The RefAtlasFrmList of all tile groups in the current atlas frame shall contain the same unique atlas frames, without, however, any restrictions in ordering of the reference atlas frames.

#### Reference atlas frame marking process

This process is invoked once per atlas frame, after decoding of an atlas tile group header and the decoding process for reference atlas frame list construction for the atlas frame as specified in clause 8.4.3.2, but prior to the decoding of the atlas tile group data. This process may result in one or more reference atlas frames in the DAFB being marked as "unused for reference" or "used for long-term reference".

A decoded atlas frame in the DAFB can be marked as "unused for reference", "used for short-term reference" or "used for long-term reference", but only one among these three at any given moment during the operation of the decoding process. Assigning one of these markings to an atlas frame implicitly removes another of these markings when applicable. When an atlas frame is referred to as being marked as "used for reference", this collectively refers to the atlas frame being marked as "used for short-term reference" or "used for long-term reference" (but not both).

Short term reference atlas frames are identified by their AtlasFrmOrderCntVal values. Long term reference atlas frames are identified by the Log2( MaxLtAtlasFrmOrderCntLsb ) least significant bits of their AtlasFrmOrderCntVal values.

For all cases, the following applies:

* For each long term reference atlas frame entry in RefAtlasFrmList, when the referred atlas frame is a short term reference atlas frame, the atlas frame is marked as "used for long-term reference".
* Each reference atlas frame in the DAFB that is not referred to by any entry in RefAtlasFrmList is marked as "unused for reference".

### Decoding process for patch units

#### Generic decoding process for patch units

Inputs to this process are the current patch index, p, and the current patch mode, atgdu\_patch\_mode[ p ].

Outputs of this process are several parameters associated with the current patch with patch index p, including its 2D and corresponding 3D location information, such as parameters Patch2dPosX[ p ], Patch2dPosY[ p ], Patch2dSizeX[ p ], Patch2dSizeY[ p ], Patch3dPosX[ p ], Patch3dPosY[ p ], Patch3dPosMinZ[ p ], PatchAxisZ[ p ], PatchProjectionFlag[ p ], PatchOrientationIndex[ p ], PatchLoDScaleX[ p ], PatchLoDScaleY[ p ], Patch45degreeProjectionRotationAxis[ p ], AuxPatch2dPosX[ p ], AuxPatch2dPosY[ p ], AuxPatch2dSizeX[ p ], AuxPatch2dSizeY[ p ], AuxPatch3dPosX[ p ], AuxPatch3dPosY[ p ], AuxPatch3dPosZ[ p ], and AuxPatchAxisZ[ p ],.

Patch2dPosX[ p ] specifies the x-coordinate of the top-left corner of the patch bounding box size for the current patch with patch index p. The value of Patch2dPosX[ p ] shall be in the range of 0 to (TileGroupWidth[ atgh\_address ] − 1), inclusive.

Patch2dPosY[ p ] specifies the y-coordinate of the top-left corner of the patch bounding box size for the current patch with patch index p. The value of Patch2dPosY[ p ] shall be in the range of 0 to (TileGroupHeight[ atgh\_address ] − 1), inclusive.

Patch2dSizeX[ p ] specifies the width of the bounding box of the current patch with patch index p. The value of Patch2dSizeX[ p ] shall be in the range of 1 to TileGroupWidth[ atgh\_address ], inclusive. It is a requirement of bitstream conformance that Patch2dPosX[ p ] + Patch2dSizeX[ p ] <= TileGroupWidth[ atgh\_address ].

Patch2dSizeY[ p ] specifies the height of the bounding box of the current patch with patch index p. The value of Patch2dSizeY[ p ] shall be in the range of 1 to TileGroupHeight[ atgh\_address ] , inclusive. It is a requirement of bitstream conformance that Patch2dPosY[ p ] + Patch2dSizeY[ p ] <= TileGroupHeight[ atgh\_address ].

Patch3dPosX[ p ] specifies the shift to be applied to the reconstructed patch points in the current patch with patch index p along the tangent axis. The value of Patch3dPosX[ p ] shall be in the range of 0 to ( 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 − 1), inclusive.

Patch3dPosY[ p ] specifies the shift to be applied to the reconstructed patch points in the current patch with patch index p along the bitangent axis. The value of Patch3dPosY[ p ] shall be in the range of 0 to ( 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 − 1), inclusive.

Patch3dPosMinZ[ p ] specifies the shift to be applied to the reconstructed patch points in the current patch with patch index p along the normal axis. The value of Patch3dPosMinZ[ p ] shall be in the range of 0 to ( 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 – 1 ), inclusive.

PatchAxisZ[ p ] is the index of the normal to the projection plane for the current patch with patch index p. The value of PatchAxisZ[ p ] shall be in range of 0 to 2, inclusive.

PatchProjectionFlag[ p ] specifies on which plane the current patch with patch index p is to be projected.

[Ed. (ATA): projection mode is a binary value and implies projection direction rather than plane?]

PatchOrientationIndex[ p ] specifies the patch orientation index to Table 9-1 for the current patch with patch index p.

PatchLoDScaleX[ p ] specifies the LOD scaling factor to be applied to the tangent axis of the current patch with patch index p.

PatchLoDScaleY[ p ] specifies the LOD scaling factor to be applied to the bi-tangent axis of the current patch with patch index p.

Patch45degreeProjectionRotationAxis[ p ]  specifies the patch projection rotation axis to be applied to the patch with index p.

[Ed. Note (JK): the limits for the auxiliary placed patches should be updated to be based on the resolution of the auxiliary video, not the resolution of the tile groups].

AuxPatch2dPosX[ p ] specifies the x-coordinate of the top-left corner of the patch bounding box size for the current patch with patch index p. The value of AuxPatch2dPosX[ p ] shall be in the range of 0 to ( TileGroupWidth[ atgh\_address ] − 1), inclusive.

AuxPatch2dPosY[ p ] specifies the y-coordinate of the top-left corner of the patch bounding box size for the current patch with patch index p. The value of AuxPatch2dPosY[ p ] shall be in the range of 0 to ( TileGroupHeight[ atgh\_address ] − 1), inclusive.

AuxPatch2dSizeX[ p ] specifies the width of the bounding box of the current patch with patch index p. The value of Patch2dSizeX[ p ] shall be in the range of 0 to TileGroupWidth[ atgh\_address ], inclusive. It is a requirement of bitstream conformance that AuxPatch2dPosX[ p ] + AuxPatch2dSizeX[ p ] <= TileGroupWidth[ atgh\_address ].

AuxPatch2dSizeY[ p ] specifies the height of the bounding box of the current patch with patch index p. The value of Patch2dSizeY[ p ] shall be in the range of 0 to TileGroupHeight[ atgh\_address ] , inclusive. It is a requirement of bitstream conformance that AuxPatch2dPosY[ p ] + AuxPatch2dSizeY[ p ] <= TileGroupHeight[ atgh\_address ].

AuxPatch3dPosX[ p ] specifies the shift to be applied to the reconstructed patch points in the current patch with patch index p along the tangent axis. The value of Patch3dPosX[ p ] shall be in the range of 0 to Min(2afps\_3d\_pos\_x\_bit\_count\_minus1 + 1, 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 − 1), inclusive.

AuxPatch3dPosY[ p ] specifies the shift to be applied to the reconstructed patch points in the current patch with patch index p along the bitangent axis. The value of Patch3dPosY[ p ] shall be in the range of 0 to Min(2afps\_3d\_pos\_y\_bit\_count\_minus1 + 1 − 1? , 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 − 1), inclusive.

AuxPatch3dPosZ[ p ] specifies the shift to be applied to the reconstructed patch points in the current patch with patch index p along the normal axis. The value of AuxPatch3dPosZ[ p ] shall be in the range of 0 to Min(2atgh\_3d\_pos\_normal\_axis\_bit\_count\_minus1 + 1 − 1?, 2gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 + 1 − 1), inclusive.

AuxPatchAxisZ[ p ] is the index of the normal to the projection plane for the current patch with patch index p. The value of PatchAxisZ[ p ] shall be in range of 0 to 2, inclusive.

[Ed. (AMT): Need a definition. (ATA) added specification]

If atgdu\_patch\_mode[ p ] is equal to I\_INTRA or P\_INTRA, then the process for decoding intra coded patches in clause 8.4.4.2 is used, with p as the input to that process and the outputs of that process are used as the output.

If atgdu\_patch\_mode[ p ] is equal to P\_SKIP, then the process for decoding skip coded patches in clause 8.4.4.3 is used, with p as the input to that process and the outputs of that process are used as the output.

If atgdu\_patch\_mode[ p ] is equal to P\_MERGE, then the process for decoding merge coded patches in clause 8.4.4.4 is used, with p as the input to that process and the outputs of that process are used as the output.

If atgdu\_patch\_mode[ p ] is equal to P\_INTER, then the process for decoding inter coded patches in clause 8.4.4.5 is used, with p as the input to that process and the outputs of that process are used as the output.

If atgdu\_patch\_mode[ p ] is equal to I\_RAW or P\_RAW, then the process for decoding RAW coded patches in clause 8.4.4.6 is used, with p as the input to that process and the outputs of that process are used as output.

If atgdu\_patch\_mode[ p ] is equal to I\_EOM or P\_EOM, then the process for decoding EOM coded patches in clause 8.4.4.7 is used, with p as the input to that process and the outputs of that process are used as output.

#### Decoding process for patch units coded in intra mode

Input to this process is the current patch index, p.

The following patch related variables are first assigned given the parsed elements in the patch data unit:

Patch2dPosX[ p ] = pdu\_2d\_pos\_x[ p ] \* PatchPackingBlockSize (8‑6)

Patch2dPosY[ p ] = pdu\_2d\_pos\_y[ p ] \* PatchPackingBlockSize (8‑7)

Patch3dPosX[ p ] = pdu\_3d\_pos\_x[ p ] (8‑8)

Patch3dPosY[ p ]= pdu\_3d\_pos\_y[ p ] (8‑9)

Patch3dPosMinZ[ p ] = Pdu3dPosMinZ[ p ] (8‑10)

PatchAxisZ[ p ] = PduProjectionPlane[ p ] % 3 (8‑11)

PatchProjectionFlag[ p ] = Clip3( 0, 1, PduProjectionPlane[ p ] / 3) (8‑12)

PatchOrientationIndex[ p ] = pdu\_orientation\_index[ p ] (8‑13)

PatchLoDScaleX[ p ] = pdu\_lod\_enabled\_flag[ p ] ? pdu\_lod\_scale\_x\_minus1[ p ] + 1: 1 (8‑14)

PatchLoDScaleY[ p ] = pdu\_lod\_enabled\_flag[ p ] ?  (pdu\_lod\_scale\_y[ p ] + (pdu\_lod\_scale\_x\_minus1[ p ] > 0) ? 1 : 2) : 1 (8‑15)

Patch45degreeProjectionRotationAxis[ p ] = Pdu45degreeProjectionRotationAxis[ p ] (8‑16)

Then the variables Patch2dSizeX[ p ] and Patch2dSizeY[ p ] are derived as follows:

Patch2dSizeX[ p ] = (pdu\_2d\_size\_x\_minus1[ p ] + 1) \* PatchSizeXQuantizer (8‑17)

Patch2dSizeY[ p ] = (pdu\_2d\_size\_y\_minus1[ p ] + 1) \* PatchSizeYQuantizer (8‑18)

Finally, the variables PatchAxisX[ p ] and PatchAxisY[ p ] are derived as follows:

– If PatchAxisZ[ p ] is equal to zero, PatchAxisX[ p ] is assigned a value of 2 and PatchAxisY[ p ] is assigned a value of 1.

– Otherwise, if PatchAxisZ[ p ] is equal to 1, PatchAxisX[ p ] is assigned a value of 2 and PatchAxisY[ p ] is assigned a value of 0.

– Otherwise (PatchAxisZ[ p ] is equal to 2) PatchAxisX[ p ] is assigned a value of 0 and PatchAxisY[ p ] is assigned a value of 1

#### Decoding process for patch units coded in skip prediction mode

Input to this process is the current patch index, p.

First, the atlas frame index, refIdx, of the first active? atlas frame in the reference atlas frame list is derived based on the process described in clause 8.4.3.2. Then the tile group with the same tile address in the first reference atlas frame as the current tile group is selected to be used as the reference for the current tile group.

If p is equal to 0, then predictorIdx is set to 0.

Then the corresponding patch index, refPatchIdx, in the refIdx atlas frame reference for the current tile group is computed as:

refPatchIdx = predictorIdx (8‑21)

and predictorIdx for the next patch is set to refPatchIdx + 1.

It is a requirement of bitstream conformance that the patch with index predictorIdx in the refIdx atlas frame reference is not a patch coded in a RAW or EOM prediction mode.

Using the variables refIdx and RefPatchIdx as inputs, the variables refPatch2dPosX, refPatch2dPosY, refPatch2dSizeX, refPatch2dSizeY, refPatch3dPosX, refPatch3dPosY, refPatch3dPosMinZ, refPatchAxisZ, refPatchProjectionFlag, refPatchAxisX, refPatchAxisY, refPatchOrientationIndex, refPatchLoDScaleX, and refPatchLoDScaleY are derived using the process described in clause 8.4.4.5.1.

Then, the associated 2D and 3D patch parameters are derived as follows:

Patch2dPosX[ p ] = refPatch2dPosX (8‑22)

Patch2dPosY[ p ] = refPatch2dPosY (8‑23)

Patch2dSizeX[ p ] = refPatch2dSizeX (8‑24)

Patch2dSizeY[ p ] = refPatch2dSizeY (8‑25)

Patch3dPosX[ p ] = refPatch3dPosX (8‑26)

Patch3dPosY[ p ] = refPatch3dPosY (8‑27)

Patch3dPosMinZ[ p ] = refPatch3dPosMinZ (8‑28)

PatchAxisZ[ p ] = refPatchAxisZ (8‑29)

PatchProjectionFlag[ p ] = refPatchProjectionFlag (8‑30)

PatchAxisX[ p ] = refPatchAxisX (8‑31)

PatchAxisY[ p ] = refPatchAxisY (8‑32)

PatchOrientationIndex[ p ] = refPatchOrientationIndex (8‑33)

PatchLoDScaleX[ p ] = refPatchLoDScaleX (8‑34)

PatchLoDScaleY[ p ] = refPatchLoDScaleY (8‑35)

Patch45degreeProjectionRotationAxis[ p ] = refPatch45degreeProjectionRotationAxis (8‑36)

[ED. AMT: the above does not currently handle EOM and RAW. To make it simpler for now I disallowed those modes although in m49464 it is mentioned that we would want to support those. More work is needed to handle these ]

#### Decoding process for patch units coded in merge prediction mode

Input to this process is the current patch index, p.

[Ed. (ATA): need to add mpdu\_override\_2d\_params\_flag[p] and mpdu\_override\_3d\_params\_flag[p] as well? Done equation numbers need to be corrected]

First, the atlas frame index, refIdx, of the mpdu\_ref\_index[ p ] + 1 ordered atlas frame in the reference atlas frame list is derived based on the process described in section 8.4.3.2. Then the tile group with the same tile address in this reference atlas frame as the current tile group is selected to be used as the reference for the current tile group.

If p is equal to 0, then predictorIdx is set to 0.

Then the corresponding patch index, refPatchIdx, in the refIdx atlas frame reference for the current tile group is computed as:

RefPatchIdx = predictorIdx (8‑37)

and predictorIdx for the next patch is set to RefPatchIdx + 1.

It is a requirement of bitstream conformance that the patch with index predictorIdx in the refIdx atlas frame reference is not a patch coded in a RAW or EOM prediction mode.

Using the variables refIdx and RefPatchIdx as inputs, the variables refPatch2dPosX, refPatch2dPosY, refPatch2dSizeX, refPatch2dSizeY, refPatch3dPosX, refPatch3dPosY, refPatch3dPosMinZ, refPatchAxisZ, refPatchProjectionFlag, refPatchAxisX, refPatchAxisY, refPatchOrientationIndex, refPatchLoDScaleX, and refPatchLoDScaleY are derived using the process described in section 8.4.4.5.1.

If mpdu\_override\_2d\_params\_flag[ p ] is equal to 1, then the associated 2D patch parameters are derived as follows:

Patch2dPosX[ p ] = refPatch2dPosX + mpdu\_2d\_pos\_x[ p ] \* PatchPackingBlockSize (8‑38)

Patch2dPosY[ p ] = refPatch2dPosY + mpdu\_2d\_pos\_y[ p ] \* PatchPackingBlockSize (8‑39)

Patch2dSizeX[ p ] = refPatch2dSizeX + mpdu\_2d\_delta\_size\_x[ p ] \* PatchSizeXQuantizer (8‑40)

Patch2dSizeY[ p ] = refPatch2dSizeY + mpdu\_2d\_delta\_size\_y[ p ] \* PatchSizeYQuantizer (8‑41)

Otherwise,

Patch2dPosX[ p ] = refPatch2dPosX (8‑42)

Patch2dPosY[ p ] = refPatch2dPosY (8‑43)

Patch2dSizeX[ p ] = refPatch2dSizeX (8‑44)

Patch2dSizeY[ p ] = refPatch2dSizeY (8‑45)

If mpdu\_override\_3d\_params\_flag[ p ] is equal to 1, then the associated 3D patch parameters are derived as follows:

Patch3dPosX[ p ] = refPatch3dPosX + mpdu\_3d\_pos\_x[ p ] (8‑46)

Patch3dPosY[ p ] = refPatch3dPosY + mpdu\_3d\_pos\_y[ p ] (8‑47)

Patch3dPosMinZ[ p ] = refPatch3dPosMinZ + Mpdu3dPosMinZ[ p ] (8‑48)

Otherwise,

Patch3dPosX[ p ] = refPatch3dPosX (8‑49)

Patch3dPosY[ p ] = refPatch3dPosY (8‑50)

Patch3dPosMinZ[ p ] = refPatch3dPosMinZ

Finally, the following variables are set:

PatchAxisZ[ p ] = refPatchAxisZ (8‑51)

PatchProjectionFlag[ p ] = refPatchProjectionFlag (8‑52)

PatchAxisX[ p ] = refPatchAxisX (8‑53)

PatchAxisY[ p ] = refPatchAxisY (8‑54)

PatchOrientationIndex[ p ] = refPatchOrientationIndex (8‑55)

PatchLoDScaleX[ p ] = refPatchLoDScaleX (8‑56)

PatchLoDScaleY[ p ] = refPatchLoDScaleY (8‑57)

Patch45degreeProjectionRotationAxis[ p ] = refPatch45degreeProjectionRotationAxis (8‑58)

#### Decoding process for patch units coded in inter prediction mode

Input to this process is the current patch index, p.

First, the atlas frame index, refIdx, of the ipdu\_ref\_index[ p ] + 1 ordered atlas frame in the reference atlas frame list is derived based on the process described in section 8.4.3.2. Then the tile group with the same tile address in this reference atlas frame as the current tile group is selected to be used as the reference for the current tile group.

If p is equal to 0, then predictorIdx is set to 0.

Then the corresponding patch index, refPatchIdx, in that atlas frame reference for the current tile group is computed as:

refPatchIdx = predictorIdx + ipdu\_patch\_index[ p ] (8‑59)

and predictorIdx for the next patch is set to refPatchIdx + 1.

It is a requirement of bitstream conformance that the patch with index predictorIdx in the refIdx atlas frame reference is not a patch coded in a RAW or EOM prediction mode.

Using the variables refIdx and refPatchIdx as inputs, the variables refPatch2dPosX, refPatch2dPosY, refPatch2dSizeX, refPatch2dSizeY, refPatch3dPosX, refPatch3dPosY, refPatch3dPosMinZ, refPatchAxisZ, refPatchProjectionFlag, refPatchAxisX, refPatchAxisY, refPatchOrientationIndex, refPatchLoDScaleX, and refPatchLoDScaleY are derived using the process described in section 8.4.4.5.1.

Then, the associated 2D and 3D patch parameters are derived as follows:

Patch2dPosX[ p ] = refPatch2dPosX + ipdu\_2d\_pos\_x[ p ] \* PatchPackingBlockSize (8‑60)

Patch2dPosY[ p ] = refPatch2dPosY + ipdu\_2d\_pos\_y[ p ] \* PatchPackingBlockSize (8‑61)

Patch2dSizeX[ p ] = refPatch2dSizeX + ipdu\_2d\_delta\_size\_x[ p ] \* PatchSizeXQuantizer (8‑62)

Patch2dSizeY[ p ] = refPatch2dSizeY + ipdu\_2d\_delta\_size\_y[ p ] \* PatchSizeYQuantizer (8‑63)

Patch3dPosX[ p ] = refPatch3dPosX + ipdu\_3d\_pos\_x[ p ] (8‑64)

Patch3dPosY[ p ] = refPatch3dPosY + ipdu\_3d\_pos\_y[ p ] (8‑65)

Patch3dPosMinZ[ p ] = refPatch3dPosMinZ + Ipdu3dPosMinZ[ p ] (8‑66)

PatchAxisZ[ p ] = refPatchAxisZ (8‑67)

PatchProjectionFlag[ p ] = refPatchProjectionFlag (8‑68)

PatchAxisX[ p ] = refPatchAxisX (8‑69)

PatchAxisY[ p ] = refPatchAxisY (8‑70)

PatchOrientationIndex[ p ] = refPatchOrientationIndex (8‑71)

PatchLoDScaleX[ p ] = refPatchLoDScaleX (8‑72)

PatchLoDScaleY[ p ] = refPatchLoDScaleY (8‑73)

Patch45degreeProjectionRotationAxis[ p ] = refPatch45degreeProjectionRotationAxis (8‑74)

##### Derivation of the inter reference patch parameters

Inputs to this process are the atlas frame reference index refIdx and patch reference index, refPatchIdx.

Outputs to this process are the variables refPatch2dPosX, refPatch2dPosY, refPatch2dSizeX, refPatch2dSizeY, refPatch3dPosX, refPatch3dPosY, refPatch3dPosMinZ, refPatchAxisZ, refPatchProjectionFlag, refPatchAxisX, refPatchAxisY, refPatchOrientationIndex, refPatchLoDScaleX, refPachLoDScaleY, and refPatch45degreeProjectionRotationAxis.

The following variables are derived, based on the patch parameters for each tile group that correspond to the same atlas tile group address in the patch reference associated with the index refIdx and set as the output of this process as follows:

refPatch2dPosX = Patch2dPosX[ refPatchIdx ] (8‑75)

refPatch2dPosY = Patch2dPosY[ refPatchIdx ] (8‑76)

refPatch2dSizeX = Patch2dSizeX[ refPatchIdx ] (8‑77)

refPatch2dSizeY = Patch2dSizeY[ refPatchIdx ] (8‑78)

refPatch3dPosX = Patch3dPosX[ refPatchIdx ] (8‑79)

refPatch3dPosY = Patch3dPosY[ refPatchIdx ] (8‑80)

refPatch3dPosMinZ = Patch3dPosMinZ[ refPatchIdx ] (8‑81)

refPatchAxisZ = PatchAxisZ[ refPatchIdx ] (8‑82)

refPatchProjectionFlag = PatchProjectionFlag[ RefPatchIdx ] (8‑83)

refPatchAxisX = PatchAxisX[ refPatchIdx ] (8‑84)

refPatchAxisY = PatchAxisY[ refPatchIdx ] (8‑85)

refPatchOrientationIndex = PatchOrientationIndex[ refPatchIdx ] (8‑86)

refPatchLoDScaleX = PatchLoDScaleX[ refPatchIdx ] (8‑87)

refPatchLoDScaleY = PatchLoDScaleY[ refPatchIdx ] (8‑88)

refPatch45degreeProjectionRotationAxis =   
 Patch45degreeProjectionRotationAxis[ refPatchIdx ] (8-48)

#### Decoding process for patch units coded in RAW mode

Input to this process is the current patch index, p.

Outputs to this process are the variables Patch3dPosX, Patch3PosY, Patch3dPosZ.

if (rpdu\_patch\_in\_auxiliary\_video\_flag [ p ] == 1 ) then the process in 8.4.4.6.1

Otherwise

##### Decoding process for patch units coded in RAW mode for patches coded in an auxiliary video

Input to this process is the current patch index, p.

Outputs to this process are the variables Patch3dPosX, Patch3PosY, Patch3dPosZ.

if (vps\_auxiliary\_video\_present\_flag[ atlasIdx ] == 1 ) then

The following patch related variables are first assigned given the parsed elements in the patch data unit:

AuxPatch2dPosX[ p ] = rpdu\_2d\_pos\_x[ p ] \* PatchPackingBlockSize (8‑89)

AuxPatch2dPosY[ p ] = rpdu\_2d\_pos\_y[ p ] \* PatchPackingBlockSize (8‑90)

Then the variables AuxPatch2dSizeX[ p ] and AuxPatch2dSizeY[ p ] are derived as follows:

AuxPatch2dSizeX[ p ] = (rpdu\_2d\_size\_x\_minus1[ p ] + 1) \* PatchSizeXQuantizer (8‑93)

AuxPatch2dSizeY[ p ] = (rpdu\_2d\_size\_y\_minus1[ p ] + 1) \* PatchSizeYQuantizer (8‑94)

AuxPatchRawPoints[ p ] = rpdu\_points\_minus1[ p ] + 1 (8‑95)

Then, the 3D patch variables AuxPatch3dPosX[ p ], AuxPatch3dPosY[ p ] and AuxPatch3dPosZ[ p ] are derived and set as output of this process as follows:

if afps\_raw\_3d\_pos\_bit\_count\_explicit\_mode\_flag is equal to 1,

AuxPatch3dPosX[ p ] = rpdu\_3d\_pos\_x[ p ] (8‑96)

AuxPatch3dPosY[ p ] = rpdu\_3d\_pos\_y[ p ] (8‑97)

AuxPatch3dPosZ[ p ] = rpdu\_3d\_pos\_z[ p ] (8‑98)

[Ed. (PA): Variables names are the same as those used in other processes except for Patch3DPosZ, specific to RAW (and different from Patch3dPosMinZ) maybe it should be defined by some sort of semantics like in 8.4.4.1: Generic decoding process for patch units. But it is specific to RAW mode I think. Ed. (PA) Suggested to rename in this clause Patch3dPosZ to Patch3dPosMinZ]

Otherwise,

raw3dLevel  = 1 << (gi\_geometry\_nominal\_2d\_bitdepth\_minus1 + 1) (8‑99)

AuxPatch3dPosX[ p ] = rpdu\_3d\_pos\_x[ p ] \* raw3dLevel (8‑100)

AuxPatch3dPosY[ p ] = rpdu\_3d\_pos\_y[ p ] \* raw3dLevel (8‑101)

AuxPatch3dPosZ[ p ] = rpdu\_3d\_pos\_z[ p ] \* raw3dLevel (8‑102)

##### Decoding process for patch units coded in RAW mode for patches coded not in an auxiliary video

The following patch related variables are first assigned given the parsed elements in the patch data unit:

Patch2dPosX[ p ] = rpdu\_2d\_pos\_x[ p ] \* PatchPackingBlockSize (8‑89)

Patch2dPosY[ p ] = rpdu\_2d\_pos\_y[ p ] \* PatchPackingBlockSize (8‑90)

Then the variables Patch2dSizeX[ p ] and Patch2dSizeY[ p ] are derived as follows:

Patch2dSizeX[ p ] = (rpdu\_2d\_size\_x\_minus1[ p ] + 1) \* PatchSizeXQuantizer (8‑93)

Patch2dSizeY[ p ] = (rpdu\_2d\_size\_y\_minus1[ p ] + 1) \* PatchSizeYQuantizer (8‑94)

PatchRawPoints[ p ] = rpdu\_points\_minus1[ p ] + 1 (8‑95)

Then, the 3D patch variables Patch3dPosX[ p ], Patch3dPosY[ p ] and Patch3dPosZ[ p ] are derived and set as output of this process as follows:

if afps\_raw\_3d\_pos\_bit\_count\_explicit\_mode\_flag is equal to 1,

Patch3dPosX[ p ] = rpdu\_3d\_pos\_x[ p ] (8‑96)

Patch3dPosY[ p ] = rpdu\_3d\_pos\_y[ p ] (8‑97)

Patch3dPosMinZ[ p ] = rpdu\_3d\_pos\_z[ p ] (8‑98)

[Ed. (PA): Variables names are the same as those used in other processes except for Patch3DPosZ, specific to RAW (and different from Patch3dPosMinZ) maybe it should be defined by some sort of semantics like in 8.4.4.1: Generic decoding process for patch units. But it is specific to RAW mode I think. Ed. (PA) Suggested to rename in this clause Patch3dPosZ to Patch3dPosMinZ]

Otherwise,

raw3dLevel  = 1 << (gi\_geometry\_nominal\_2d\_bitdepth\_minus1 + 1) (8‑99)

Patch3dPosX[ p ] = rpdu\_3d\_pos\_x[ p ] \* raw3dLevel (8‑100)

Patch3dPosY[ p ] = rpdu\_3d\_pos\_y[ p ] \* raw3dLevel (8‑101)

Patch3dPosMinZ[ p ] = rpdu\_3d\_pos\_z[ p ] \* raw3dLevel (8‑102)

#### Decoding process for patch units coded in EOM mode

Input to this process is the current EOM patch index, p.

Outputs to this process are the variables EomPatch2dPosX, EomPatch2dPosY, EomPatch2dSizeX, and EomPatch2dSizeY.

The following patch related variables are first assigned given the parsed elements in the patch data unit:

EomPatch2dPosX[ p ] = epdu\_2d\_pos\_x[ p ] \* PatchPackingBlockSize (8‑103)

EomPatch2dPosY[ p ] = epdu\_2d\_pos\_y[ p ] \* PatchPackingBlockSize (8‑104)

Then the variables EomPatch2dSizeX[ p ] and EomPatch2dSizeY[ p ] are derived as follows:

EomPatch2dSizeX[ p ] = (epdu\_2d\_size\_x\_minus1[ p ] + 1) \* PatchSizeXQuantizer (8‑107)

EomPatch2dSizeY[ p ] = (epdu\_2d\_size\_y\_minus1[ p ] + 1) \* PatchSizeYQuantizer (8‑108)

[Ed (PA): The process is described for one EOM patch only... Dependency between geometry patches and EOM attribute points make it difficult to generalize. It is proposed to have a contribution proposing either to have 1 EOM attribute patch per associated geometry patch or 1 EOM attribute patch for all associated geometry patches or signal the first geometry patch associated to the first EOM attribute patch. Alternatively, all EOM attribute patches should be decoded and information is referenced by geometry patch index u0, v0, sizeU0, sizeV0, EomPointOffset. Ed. (PA): This should be fixed with m51037 adopted in Geneva meeting (MPEG128)] [Ed. (PA): The issue seems address since latest adoptions]

### Decoding process of the block to patch map

Inputs to this process are:

* the current atlas tile group address, tgAddress
* the total number of patches in the current atlas tile group, PfduTotalNumberOfPatches
* the Patch2dPosX, Patch2dPosY, Patch2dSizeX, Patch2dSizeY arrays
* the PatchPackingBlockSize, asps\_frame\_height, and asps\_frame\_width elements of the current active ASPS

Outputs of this process are a two-dimensional array BlockToPatchMap and its width, BlockToPatchMapWidth, and height, BlockToPatchMapHeight, where:

BlockToPatchMapWidth = Ceil( TileGroupWidth[tgAddress ] ÷ PatchPackingBlockSize )

BlockToPatchMapHeight = Ceil( TileGroupHeight[tgAddress ] ÷ PatchPackingBlockSize )

All elements of BlockToPatchMap are first initialized to −1 as follows:

for( y = 0; y < BlockToPatchMapHeight; y++ )

for( x = 0; x < BlockToPatchMapWidth; x++ )

BlockToPatchMap[ y ][ x ] = −1

Then the BlockToPatchMap array is updated as follows:

for( patchIdx = 0; patchIdx < PfduTotalNumberOfPatches; patchIdx++ ) {

mode = atgdu\_patch\_mode[ patchIdx ]

if ((( atgh\_type == I\_TILE\_GRP ) && ( mode != I\_RAW ) && ( mode != I\_EOM )) ||  
 (( atgh\_type == P\_TILE\_GRP ) && (mode != P\_RAW) && ( mode != P\_EOM ) ||   
 ( atgh\_type == SKIP\_TILE\_GRP )) {

xOrg = Patch2dPosX[ patchIdx ] / PatchPackingBlockSize

yOrg = Patch2dPosY[ patchIdx ] / PatchPackingBlockSize

for( y = 0; y < Patch2dSizeX [ patchIdx ]/ PatchPackingBlockSize ; y++)

for( x = 0; x < Patch2dSizeY[ patchIdx ] / PatchPackingBlockSize ;x++) {

if(( asps\_patch\_precedence\_order\_flag == 0 )  ||  
 ( BlockToPatchMap[ yOrg + y ][ xOrg + x ] == −1))   
 BlockToPatchMap[ yOrg + y ][ xOrg + x ] = patchIdx

}

}

}

}

[Ed. (AMT): need to specify the TileGroupWidth/Height lists. Also it should be noted that now, for tile groups, the size of the Map is tile group based, not full frame based. It should be checked if this has any impact.]

[Ed. (AMT): It should be stated that no other patch should intersect a RAW patch, although it is okay for a RAW patch to intersect another patch. Otherwise that could have an impact on RAW point decoding. Furthermore,

[Ed. (ATA): Do we permit to use different partitions for referencing?][Ed. (AMT) This is prohibited.]

### Conversion of tile group level information to Atlas level information

In this clause, the patches from different tile groups are combined into a single list. Let the total number of such patches in the current Atlas, for the current point cloud frame be AtlasTotalNumberOfPatches. The array AtlasPatchMode[ p ] stores the patch mode for the patch with index p, p = 0..( AtlasTotalNumberOfPatches – 1), in the combined list of patches. The BlockToPatch arrays for the tile groups are also combined and each entry points to a patch from the combined list of patches. Patch2dPosX and Patch2dPosY are with respect to the origin of the current atlas instead of the current tile group. The arrays Patch2dSizeX, Patch2dSizeY, Patch3dPosX, Patch3dPosY, and Patch3dPosZ corresponding to the tile groups are also combined to form arrays at the atlas level.

[Ed. (RJ): Describe more precisely by looping over tile groups.]

## Occupancy map video decoding process

A video decoding process is invoked using the occupancy video bitstream and its associated codec, specified by occ\_occupancy\_codec\_id, as inputs. Outputs of this process are the frame rate of the occupancy map video, OccFrameRate, the decoded and output ordered occupancy video frames, OccFrame[ orderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, OccBitdepth[ orderIdx ], width, OccWidth[ orderIdx ], and height, OccHeight [ orderIdx ], where orderIdx is the output order index of the decoded occupancy map video frames, compIdx corresponds to the colour component index and is equal to 0, y is the row index in the decoded frame and is in the range of 0 to OccHeight[ orderIdx ] − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to OccWidth[ orderIdx ] − 1, inclusive.

[Ed. (RJ): Is it agreed that occupancy map video is always at the nominal frame rate? In that case, OccFrameRate is not needed to be output.]

If asps\_eom\_patch\_enabled\_flag is equal to 0, each decoded occupancy video frame with output order index of orderIdx, compIdx of 0, is thresholded as follows:

The variable lossyOccMapThreshold is derived as follows:

lossyOccMapThreshold[ orderIdx ] =   
 oi\_lossy\_occupancy\_map\_compression\_threshold << ( OccBitdepth[ orderIdx ] – 8 ).

for( y=0; y < OccHeight[ orderIdx ]; y++ )  
 for( x=0; x < OccWidth[ orderIdx ]; x++ )  
 if( OccFrame[ orderIdx ][ compIdx ][ y ][ x ] <= lossyOccMapThreshold[ orderIdx ] )  
 OccFrame[ orderIdx ][ compIdx ][ y ][ x ] = 0  
 else  
 OccFrame[ orderIdx ][ compIdx ][ y ][ x ] = 1

The decoded (possibly thresholded) occupancy map video frames shall be upscaled to match the nominal resolution of the decoded video frames, namely asps\_frame\_width pixels wide and asps\_frame\_height pixels high. The occupancy map upscaling process is defined in TBD.

[Ed. (AMT): Above sentence should be moved into section 9]

NOTE – Any existing video codec such as AVC or HEVC or any future defined video codec may be used if included in occ\_occupancy\_codec\_id.

[Ed. (JL): the note about video coding appears in several places, it’d be better to add a generic comment somewhere else but I have no concrete proposal so far.]

[Ed. (AMT): The same comment about "standardized" applies. Also, this should likely also have the resolution of the video, not the nominal one.]

[Ed. (ATA): time dependency of lossyOccMapThreshold[ orderIdx ] – what is the use case?]

# Reconstruction process

## General

[Ed. (All): Add diagram illustrating the complete point cloud decoding and reconstruction process in section just before Section 8 Decoding.]

Input to this process is a set of decoded video sub-bitstreams, corresponding to the occupancy, geometry, and if available, attribute information of the point cloud sequence, as well as the decoded atlas information. The process also takes as inputs the syntax elements and upper-case variables from clause 7 and 8. The decoded video sub-bitstreams and the decoded atlas information is referred to as a decoded point cloud frame.

[Ed. (RJ): Perhaps this should be defined in definitions section as well as in defining the output of the decoding process in section 8.]

The output of this process is a sequence of reconstructed, possibly smoothed, point cloud frames. A reconstructed point cloud frame consists of a set of points with (x, y, z) coordinates. Optionally, it may also contain one or more sets of attributes, as shown in Table 7‑2, that are associated with each point.

[Ed. (RJ): Should "that shall be" be replaced by "that are"? Done.]

[Ed. (RJ): Should a definition of reconstructed point cloud frame be added in the definitions section?]

For the sequence of decoded point cloud frames, the following ordered steps apply

* Chroma format, resolution, and frame rate conversions are performed, if required, as specified in clause 9.2.
* Point cloud reconstruction is performed as specified in clause 9.3.
* Point cloud smoothing is performed as specified in clause 9.6.

[Ed. (All): Should point cloud reconstruction be invoked before upscaling in certain cases such as pixel\_deinterleaving? Or perhaps upscaling could be restricted.][Ed. (RJ): Resolved. Always do upscaling first. The specification does not break in the above case, only that the performance may not be optimal and is untested. My preference would be to put a normative restriction or add a note. There are many such cases where different combination of tools is untested or is not allowed by the test model encoder.]

[Ed. (SS): Should smoothing be optional?][Ed. (RJ): The smoothing section will describe its dependence on certain flags. Perhaps, we should add preambles to the ordered states above clarifying when those are applied.]

[Ed. (AMT): Now the process is specified using tile groups, not frames. We need to check if anything needs updating because of this. Response: If decoding in section 8 already places patches according to the atlas coordinates (not wrt tile-group coordinates), we are all set.]

## Chroma format, resolution, and frame rate conversion

[Ed. (All): Strict conformance to point B can be defined using the above. However using a different upsampling process may still be considered conforming.]

Inputs to this process are:

* the decoded geometry frames, GeoFrame, and their associated bitdepth, GeoBitdepth, width, GeoWidth, height, GeoHeight, and frame rate GeoFrameRate
* the decoded attribute frames, AttrFrame, and their associated bitdepth, AttrBitdepth width, AttrWidth, height, AttrHeight, and frame rate AttrFrameRate, if available
* the decoded occupancy frames, OccFrame, and their associated bitdepth, OccBitdepth, width, OccWidth, height, OccHeight, and frame rate OccFrameRate
* the BlockToPatchMap arrays and their associated width, BlockToPatchMapWidth, height, BlockToPatchMapHeight, and frame rate BlockToPatchMapRate
* the decoded atlas tile groups and their associated patch information, Patch2dPosX, Patch2dPosY, Patch2dSizeX, Patch2dSizeY, Patch3dPosX, Patch3dPosY, Patch3dPosMinZ, PatchAxisZ, PatchOrientationIndex, PatchLoDScaleX, PatchLoDScaleY, and PatchProjectionFlag
* the atlas with index atlasIdx, width, asps\_frame\_width, height, asps\_frame\_height, frame rate, vpcc\_frame\_rate, and geometry bitdepth, 2gi\_geometry\_2d\_log2bitdepth

[Ed. (RJ): Clause 8.4.6 already converts this information to frame level. So that is what should be input over here.]

Outputs of this process are:

* the upscaled to the atlas width, asps\_frame\_width, atlas height, asps\_frame\_height, nominal frame rate, vpcc\_frame\_rate, and geometry bitdepth, gi\_geometry\_2d\_nominal\_bitdepth\_minus1, geometry frames, GeoFrameNR.
* the upscaled to the atlas width, asps\_frame\_width, height, asps\_frame\_height, and frame rate, vpcc\_frame\_rate, and attribute bitdepth, ai\_attribute\_nominal\_2d\_bitdepth\_minus1, attribute frames, AttrFrameNR, occupancy frames, OccFrameNR, and block to patch map arrays, BlockToPatchMap.
* the upscaled to the atlas width, asps\_frame\_width, height, asps\_frame\_height, and frame rate, vpcc\_frame\_rate, atlas frames and their associated patch information, Patch2dPosX, Patch2dPosY, Patch2dSizeX, Patch2dSizeY, Patch3dPosX, Patch3dPosY, Patch3dPosMinZ, PatchAxisZ, PatchOrientationIndex, PatchLoDScaleX, PatchLoDScaleY, and PatchProjectionMod.

The upscaling process is outside the scope of this document. At the end of this process, all such information should be in sync and time-aligned.

[Ed (AMT): Should we assume that patches are already at the nominal frame rate? That would of course impact the BlockToPatchMap and the patch information since no frame rate conversion would be required for those. There is also a question if frame rate should be kept at the nominal frame rate for the Occupancy information. We should also not exclude separately coded RAW information videos.]

[Ed (AMT): upscaling of the atlas frames might not currently make much sense. If we restrict also that this is always at the nominal resolution/frame-rate then such is not needed. Note that I did not also change the variable names to avoid having to do the editing work and reverting it back.]

[Ed (AMT): Bitdepth is included also for the occupancy since maybe we would want to support "lossy" encoding + quantization for its signaling. Even if that is not added, it does not hurt including this here.]

[Ed (AMT): we need to define nominal frame rate and geometry bitdepth in the VPS. How should we also specify frame rate? Maybe use the same mechanism as in HEVC, with the avg\_pic\_rate[ i ] concept, i.e. avg\_pic\_rate[ i ] indicates the average picture rate, in units of picture per 256 seconds ]

[Ed (RJ): Add process for upscaling occupancy map, geometry, and attribute videos from decoded to nominal resolution.]

[Ed (ATA): To make it more explicit, should we use vps\_avg\_frame\_rate!]

[Ed (ATA): Should the above inputs/ouputs be indexed by atlasId or use current atlas, instead?]

## Point cloud reconstruction

For the current point cloud frame with index OrderIdx in display order, and the current atlas with index AtlasIdx, the inputs to this process are:

* GeoFrameNR[ mapIdx ][ OrderIdx ][ 0 ][ y ][ x ], the 0th component of the decoded geometry frame with index OrderIdx at nominal resolution (after upscaling if necessary), denoted by GFrame[ mapIdx ][ y ][ x ],
* AttrFrame[ attrIdx ][ mapIdx ][ OrderIdx ][ compIdx ][ y ][ x ], the decoded attribute frame with index OrderIdx at nominal resolution (after upscaling, if necessary), denoted by AFrame[ attrIdx ][ mapIdx ][ compIdx ][ y ][ x ], and
* OccFrame[ OrderIdx ][ 0 ][ y ][ x ], the decoded occupancy map frame with index OrderIdx at nominal resolution (after upscaling, if necessary), denoted by OFrame[ y ][ x ].

The outputs of this process are:

* PointCnt, the number of points in the reconstructed point cloud frame.
* RecPcGeo[ n ][ k ], a container holding a list of coordinates of the points in the reconstructed point cloud frame, where n = 0 .. PointCnt – 1, k = 0 .. 2, and
* when ai\_attribute\_count[ AtlasIdx ] is greater than 0, RecPcAttr[ n ][ attrIdx ][ compIdx ], a container holding a list of attributes corresponding to the points in the reconstructed point cloud frame, where n = 0 .. PointCnt – 1, attrIdx = 0 .. ai\_attribute\_count – 1, and compIdx = 0 .. ai\_attribute\_dimension\_minus1[ attrIdx ].

[Ed. (PA): Should we add as outputs new PointToPact and AttrPresent?]

The following applies:

* A variable PointCnt is initialized to 0. Arrays RecPcGeo, RecPcAttr, PointToPatch, and AttrPresent are initially empty.
* For p = 0 .. AtlasTotalNumberOfPatches − 1, the following applies:
  + - If AtlasPatchMode[ p ] is I\_RAW or P\_RAW, clause 9.5 is invoked with variable PointCnt, arrays GFrame[ 0 ] and AFrame, and patch index p as inputs. The outputs are the updated variables PointCnt, RecPcGeo, RecPcAttr, PointToPatch, and AttrPresent.
    - Otherwise, if AtlasPatchMode[ p ] is I\_EOM or P\_EOM, clause 9.4.4 is invoked with OFrame, GFrame, AFrame, PatchProjectionFlag, PatchPackingBlockSize and patch index p as inputs. The outputs are updated PointCnt, RecPcGeo, RecPcAttr, PointToPatch, and AttrPresent. [Ed. (RJ, PA, ATA): It is proposed that the geometry of the points is extracted from the occupancy map corresponding regular patch in clause 9.4.4 as well.]
    - Otherwise (AtlasPatchMode[ p ] is I\_INTRA, P\_SKIP, P\_MERGE, P\_INTRA, or P\_INTER), the following applies:
      * For u = 0..Patch2dSizeX[ p ] − 1, v = 0..Patch2dSizeY[ p ] – 1, the following applies:
        + The conversion of patch coordinates to atlas coordinates as specified in clause 9.4.5 is invoked with patch coordinates ( u, v ), and patch index p, as inputs and atlas coordinates ( x, y ) as output.
        + The variables subBlkX and subBlkY are calculated as follows:

subBlkY = y / PatchPackingBlockSize   
 subBlkX = x / PatchPackingBlockSize

* + - * + If OFrame[ y ][ x ] is not equal to 0 and AtlasBlockToPatchMap[ subBlkY ][ subBlkX ] is equal to p, the following applies: [Ed. (RJ): It is assumed that precedenceFlag has already been taken into account when BlockToPatchMap is derived. No action needed.]

If asps\_pixel\_deinterleaving\_enabled\_flag is equal to 1, the following applies:

For mapIdx = 0 .. asps\_map\_count\_minus1, Clause 9.4.1 is invoked [Ed. (RJ): Specify inputs and outputs].

Otherwise, if asps\_point\_local\_reconstruction\_enabled\_flag is equal to 1, the following applies:

For mapIdx = 0 .. asps\_map\_count\_minus1, Clause 9.4.2 is invoked [Ed. (RJ): Specify inputs and outputs].

Otherwise ( asps\_pixel\_deinterleaving\_enabled\_flag is equal to 0 and asps\_point\_local\_reconstruction\_enabled\_flag is equal to 0 ), clause 9.4.3 is invoked [Ed. (RJ): Specify inputs and outputs].

If asps\_eom\_patch\_enabled\_flag is equal to 1, clause 9.4.4 is invoked with GFrame, OFrame, AFrame, patch index p, PatchProjectionFlag[ p ], and PatchPackingBlockSize as inputs and the outputs are updated PointCnt, RecPcGeo, RecPcAttr, PointToPatch, and AttrPresent. [Ed. (RJ) (PA) (ATA): It is suggested that this is done in I\_EOM or P\_EOM patch processing.]

* The points for which no attributes are found in the sub-bitstream, are assigned attributes as follows:
* For n = 0 .. ( PointCnt – 1 ), the following applies:
  + If Patch45degreeProjectionRotationAxis[ PointToPatch[ n ] ] is greater than 0, clause 9.4.7 is invoked with patch index p and the three-dimensional vector RecPcGeo[ n ] as inputs. The output of this process is an updated vector RecPcGeo[ n ].
* If AttrPresent[ n ] is equal to 0, the following applies:
  + A point from RecPcGeo that is nearest to RecPcGeo[ n ] and which satisfies the following conditions is determined. Let the index of such a point in RecPcGeo be m.

AttrPresent[ m ] = 1, and  
 PointToPatch[ m ] = = PointToPatch[ n ].

* + The attributes corresponding to the point from RecPcGeo with index m are copied as follows:

for( a = 0; a < ai\_attribute\_count[ atlasIdx ]; a++ )  
 for( c = 0; c < ai\_attribute\_dimension\_minus1[ atlasIdx ][ a ]; c++ )  
 RecPcAttr[ n ][ a ][ c ] =RecPcAttr[ m ][ a ][ c ]

[Ed. (RJ) (PA) (ATA): Align text to software.]

[Ed. (RJ): Need to add proper inputs and outputs for the invocation above.]

[Ed. (PA): At some point we need adding references to smoothing processes or maybe in an annex as being currently under discussions. Indeed 9.6 Smoothing process are unreferenced yet in the reconstruction process.]

## Reconstruction of points for non-RAW intra and inter coded patches

### Reconstruction of points for non-RAW intra and inter coded patches when asps\_pixel\_deinterleaving\_enabled\_flag is equal to 1

[Ed. (RJ): Add figure showing pixel interleaving.]

[Ed. (SS): Add note that only two layers are supported in V-PCC for pixel interleaving.][Ed. (RJ): Resolved. According to Gothenburg decision, the text should be general to handle any number of layers. The restriction is introduced in A.4.1.]

[Ed. (PA): Add inputs and outputs of the process.]

* + The 3D coordinates corresponding to the point with patch coordinates ( u, v ) are derived by invoking clause 9.4.6 with ( u, v ), patch index pIdx, and GFrame[ 0 ][ y ][ x ] as inputs. The output is a three-dimensional vector, point3D, containing the ( x, y, z ) coordinates of the point.
* The vector, point3D, is copied over to a new entry in RecPcGeo.

for( n=0; n < 3; n++ ) {  
 RecPcGeo[ PointCnt ][ n ] = point3D[ n ]  
}  
PointToPatch[ PointCnt ] = pIdx

* For aIdx = 0 .. ai\_attribute\_count[ AtlasIdx ]− 1, attributes are assigned to the newly added point as follows:

attrDim = ai\_attribute\_dimension\_minus1[ AtlasIdx ][ aIdx ] + 1  
for( cIdx = 0; cIdx  <  attrDim; cIdx+ + )  
 RecPcAttr[ PointCnt ][ aIdx ][ cIdx ] = AFrame[ aIdx ][ mapIdx ][ cIdx ][ y ][ x ]  
AttrPresent[ PointCnt ] = 1

* The variable PointCnt is incremented by 1.
* If asps\_pixel\_deinterleaving\_map\_flag[ mapIdx ] is equal to 1, the following applies:
  + The neighbours of the pixel, which all belong to the patch with index pIdx, are identified and their minimum and maximum depth values are derived as follows:
    - The variable validNeighbourCount and and array depthNeighbours are initialized as follows:

validNeighbourCount = 0  
for( k = 0; k < 4; k++ )  
 depthNeighbours[ k ] = 0

* + - The neighbours of the pixel that belong to the patch with index pIdx, are identified and their minimum and maximum geometry values are derived as follows:

minDepth = point3D[ PatchAxisZ[ pIdx ] ]  
maxDepth = point3D[ PatchAxisZ[ pIdx ] ].

[Ed. Note: the above seems incorrect since it is exceeding the array range of RecPcGeo. It is expected that this should be using the current point's depth value][Ed. (RJ): Fixed.]

if( x > 0) {  
 subBlkY = y / PatchPackingBlockSize   
 subBlkX = (x – 1) / PatchPackingBlockSize   
 if( ( OFrame[ y ][ x – 1 ] != 0 ) && (BlockToPatchMap[ subBlkY ][ subBlkX ] == pIdx ) ) {  
 depthNeighbours[ 0 ] = Patch3dPosMinZ[ pIdx ] +  
 (1 – 2 \* PatchProjectionFlag[ pIdx ] ) \* GFrame[ 0 ][ y ][ x – 1 ]  
 validNeighbourCount++  
 minDepth = Min( minDepth, depthNeighbour[ 0 ] )  
 maxDepth = Max( maxDepth, depthNeighbour[ 0 ] )  
 }  
}

if( x < ( asps\_frame\_width – 1 ) ) {  
 subBlkY = y / PatchPackingBlockSize   
 subBlkX = (x + 1) / PatchPackingBlockSize  
 if( ( OFrame[ y ][ x + 1 ] != 0 ) && (BlockToPatchMap[ subBlkY ][ subBlkX ] == pIdx)) {  
 depthNeighbours[ 1 ] = Patch3dPosMinZ [ pIdx ] +   
 (1 – 2 \* PatchProjectionFlag[ pIdx ] ) \* GFrame[ 0 ][ y ][ x + 1 ]  
 validNeighbourCount++  
 minDepth = Min( minDepth, depthNeighbour[ 1 ] )  
 maxDepth = Max( maxDepth, depthNeighbour[ 1 ] )  
 }  
}

if( y > 0 ) {  
 subBlkY = (y − 1) / PatchPackingBlockSize   
 subBlkX = x / PatchPackingBlockSize

if( ( OFrame[ y− 1 ][ x ] != 0 ) && ( BlockToPatchMap[ subBlkY ][ subBlkX ] == pIdx ) ) { depthNeighbours[ 2 ] = Patch3dPosMinZ [ pIdx ] +   
 (1 – 2\* PatchProjectionFlag[ pIdx ]) \* GFrame[ 0 ][ y – 1 ][ x ]  
 validNeighbourCount++  
 minDepth = Min( minDepth, depthNeighbour[ 2 ] )  
 maxDepth = Max( maxDepth, depthNeighbour[ 2 ] )  
 }  
}

if( y < ( asps\_frame\_height – 1 ) ) {

subBlkY = (y + 1) / PatchPackingBlockSize   
 subBlkX = x / PatchPackingBlockSize

if( ( OFrame[ y+ 1 ][ x ] != 0 ) && (BlockToPatchMap[ subBlkY ][ subBlkX ] == pIdx ) ){   
 depthNeighbours[ 3 ] = Patch3dPosMinZ [ pIdx ] +   
 (1 – 2 \* PatchProjectionFlag[ pIdx ]) \*  GFrame[ 0 ][ y + 1 ][ x ]  
 validNeighbourCount++  
 minDepth = Min( minDepth, depthNeighbour[ 3 ] )  
 maxDepth = Max( maxDepth, depthNeighbour[ 3 ] )  
 }  
}

* If(( x + y ) % 2 ) is equal to 1 the following applies:
* The missing depth value is estimated as follows:

depth1 = point3D[ PatchAxisZ[ pIdx ] ]  
for( k = 0; k < 3; k++ )  
 pos3D[ k ] = point3D[ k ]  
if( PatchProjectionFlag[ pIdx ] = = 0 )  
 depth0 = Clip3(depth1 – asps\_surface\_thickness\_minus1 – 1, depth1, minDepth )   
else  
 depth0 = Clip3( depth1, depth1 + asp\_surface\_thickness\_minus1 + 1, maxDepth )

* If depth0 is not equal to depth1, the following applies:

pos3D[ PatchAxisZ[ pIdx ] ] = depth0  
for( k = 0; k < 3; k++ )  
 RecPcGeo[ PointCnt ][ k ] = pos3D[ k ]  
AttrPresent[ PointCnt ] = 0  
PointToPatch[ PointCnt ] = pIdx  
PointCnt = PointCnt + 1

* Otherwise ( ( x + y ) % 2 ) is equal to 0 ) the following applies:
* The missing depth value is estimated as follows:

depth0 = point3D[ PatchAxisZ[ pIdx ] ]  
for( k = 0; k < 3; k++ )  
 pos3D[ k ] = point3D[ k ]  
averageDepth = 0  
for( k = 0; k < 4; k++ )  
 averageDepth += depthNeighbours[ k ]  
averageDepth = averageDepth / ValidNeighbourCount  
if( PatchProjectionFlag[ pIdx ] = = 0 )  
 depth1 = Clip3( depth0,  depth0 + asps\_surface\_thickness\_minus1 + 1, averageDepth )  
else  
 depth1 = Clip3( depth0 − asps\_surface\_thickness\_minus1 − 1, depth0, averageDepth )

* If depth1 is not equal to depth0, the following applies :

pos3D[ PatchAxisZ[ pIdx ] ] = depth1  
for( k = 0; k < 3; k++ )  
 RecPcGeo[ PointCnt ][ k ] = pos3D[ k ]  
AttrPresent[ PointCnt ] = 0  
PointToPatch[ PointCnt ] = pIdx  
PointCnt = PointCnt + 1

* The points between depth0 and depth1 are filled as follows:

maxDepth = Max( depth0, depth1 )  
 minDepth = Min( depth0, depth1 )

* For s = 1..( maxDepth – minDepth – 1), the following applies:

pos3D[ PatchAxisZ[ pIdx ] ] = midDepth + s  
for( k = 0; k < 3; k++ )  
 RecPcGeo[ PointCnt ][ k ] = pos3D[ k ]  
AttrPresent[ PointCnt ] = 0  
PointToPatch[ PointCnt ] = pIdx  
PointCnt = PointCnt + 1

### Reconstruction of points for non-RAW intra and inter coded patches when plri\_point\_local\_reconstruction\_map\_flag is equal to 1

Inputs to this process are:

* PointCnt, the number of points in the reconstructed point cloud frame.
* GFrame[ mapIdx ][ y ][ x ], where mapIdx = 0..asps\_map\_count\_minus1, y = 0 .. asps\_frame\_height – 1, and x = 0 .. asps\_frame\_width – 1.
* OFrame[ y ][x], where y = 0 .. asps\_frame\_height – 1 and x = 0 .. asps\_frame\_width – 1.
* AFrame[ aIdx ][ mapIdx ][ cIdx ][ y ][ x ], where mapIdx = 0 .. asps\_map\_count\_minus1, y = 0 .. asps\_frame\_height – 1, x = 0 .. asps\_frame\_width – 1, aIdx = 0 .. ai\_attribute\_count – 1, and cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ].
* the index of the patch, pIdx.
* PatchProjectionFlag[ pIdx ], the plane on which the current patch with patch index pIdx is projected
* PatchPackingBlockSize, the size of the block used for the horizontal and vertical placement of the patches within the atlas.
* asps\_point\_local\_reconstruction\_information( mapIdx ) decoded syntax elements of clause 7.3.6.2, where mapIdx = 0 .. asps\_map\_count\_minus1.
* point\_local\_reconstruction\_data( pIdx ) decoded syntax elements of clause 7.3.7.8.

Outputs to this process are:

* recPcGeo[ n ][ k ], a container holding a list of coordinates of the points in the reconstructed point cloud frame, where n = 0 .. PointCnt – 1 and k = 0 .. 2.
* recPcAttr[ n ][ aIdx ][ cIdx ], a container holding a list of attributes corresponding to the points in the reconstructed point cloud frame, where n = 0 .. PointCnt – 1, aIdx = 0 .. ai\_attribute\_count – 1, and cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ].

When asps\_point\_local\_reconstruction\_enabled\_flag equal 1, the process described in this clause is used to reconstruct the point cloud. The point local reconstruction process is applied to each occupied pixel of the geometry frame. As an output, the point cloud is updated with new points.

The following processes apply for each map of index mapIdx with mapIdx = 0..asps\_map\_count\_minus1.

Let log2BlockSize = log2(PatchPackingBlockSize).

When plri\_point\_local\_reconstruction\_map\_flag[ mapIdx ] is equal to 1, the variable plrMode[ bIdxY ][ bIdxX ] is derived in a first step for each block of coordinates (bIdxX, bIdxY) for the map mapIdx of the patch pIdx, as follows:

if( plrd\_level[ mapIdx ][ pIdx ] = = 0 ) {

for( bIdxY = 0; bIdxY <= ( ( Patch2dSizeY[ pIdx ] − 1 )>> log2BlockSize ); bIdxY++ )

for( bIdxX = 0; bIdxX <= ( ( Patch2dSizeX[ pIdx ] − 1 )>> log2BlockSize ); bIdxX++ ) {

plrMode[ bIdxY ][ bIdxX ] = 0

if( plrd\_present\_block\_flag[ mapIdx ][ pIdx ][ blockCnt ] )

plrMode[ bIdxY ][ bIdxX ] = plrd\_block\_mode\_minus1[ mapIdx ][ pIdx ][ blockCnt ] + 1

blockCnt ++

}

} else {

for( bIdxY = 0; bIdxY <= ( ( Patch2dSizeY[ pIdx ] − 1 )>> log2BlockSize ); bIdxY++ )

for( bIdxX = 0; bIdxX <= ( ( Patch2dSizeX[ pIdx ] − 1 )>> log2BlockSize ); bIdxX++ ) {

plrMode[ bIdxY ][ bIdxX ] = 0

if( plrd\_present\_flag[ mapIdx ][ pIdx ] )

plrMode[ bIdxY ][ bIdxX ] = plrd\_mode\_minus1[ mapIdx ][ pIdx ] + 1

}

}

Then, in a second step, the variables plrInterpolateFlag[ bIdxY ][ bIdxX ], plrFillingFlag[ bIdxY ][ bIdxX ], plrMinimumDepth[ bIdxY ][ bIdxX ] and plrNeighbour[ bIdxY ][ bIdxX ], are derived as follows:

for( bIdxY = 0; bIdxY <= ( ( Patch2dSizeY[ pIdx ] − 1 )>> log2BlockSize ); bIdxY++ )

for( bIdxX = 0; bIdxX <= ( ( Patch2dSizeX[ pIdx ] − 1 )>> log2BlockSize ); bIdxX++ ) {

plrInterpolateFlag[ bIdxY ][ bIdxX ] = plri\_interpolate\_flag[ mapIdx ][ plrMode[ bIdxY ][ bIdxX ] ]

plrFillingFlag[ bIdxY ][ bIdxX ] = plri\_filling\_flag[ mapIdx ][ plrMode[ bIdxY ][ bIdxX ] ]

plrMinimumDepth[ bIdxY ][ bIdxX ] = plri\_minimum\_depth[ mapIdx ][ plrMode[ bIdxY ][ bIdxX ] ] plrNeighbour[ bIdxY ][ bIdxX ] = plri\_neighbour\_minus1[ mapIdx ][ plrMode[ bIdxY ][ bIdxX ] ] + 1

}

The third step consists of the point local reconstruction process which applies as follows:

For u = 0..Patch2dSizeX[ pIdx ] − 1, v = 0..Patch2dSizeY[ pIdx ] – 1, the following applies:

* The conversion of patch coordinates to atlas coordinates as specified in clause 9.4.5 is invoked with patch coordinates (u, v) and patch index pIdx as inputs and atlas coordinates (x, y) as output.

– If OFrame[ y ][ x ] is not equal to 0, the following steps apply:

* The variable deltaDepth is initialized as follows:

deltaDepth = 0

* If plrInterpolateFlag[ y >> log2BlockSize ][ x >> log2BlockSize ] is equal to 1, then
  + If PatchProjectionFlag[ pIdx ] is equal to 0 then,
    - Find the neighbouring pixel of coordinates (xn, yn) in the geometry frame belonging to the current patch with the largest geometry value such that GFrame[ mapIdx ][ yn ][ xn ] – GFrame[ mapIdx ][ y ][ x ] is equal or smaller than (asps\_surface\_thickness\_minus1 + 1).

The neighbourhood is defined as a square window of size 2 \* plrNeighbour[ y >> log2BlockSize ][ x >> log2BlockSize ] + 1 centered around (x, y).

* + - Set deltaDepth as follows:

deltaDepth = GFrame[ mapIdx ][ yn ][ xn ] – GFrame[ mapIdx ][ y ][ x ] – 1

* + If PatchProjectionFlag[ pIdx ] is equal to 1 then,
    - Find the neighbouring pixel of coordinates (xn, yn) in the geometry frame belonging to the current patch with the smallest geometry value such that GFrame[ mapIdx ][ yn ][ xn ] – GFrame[ mapIx ][ y ][ x ] is greater than or equal to – (asps\_surface\_thickness\_minus1 + 1).

The neighbourhood is defined as a square window of size 2 \* plrNeighbour[ y >> log2BlockSize ][ x >> log2BlockSize ] + 1 centered around (x,y).

* + - Set deltaDepth as follows:

deltaDepth = GFrame[ mapIdx ][ yn ][ xn ] – GFrame[ mapIdx ][ y ][ x ] – 1

* Update deltaDepth as follows:
  + If PatchProjectionFlag[ pIdx ] is equal to 0 then,

deltaDepth = Max( deltaDepth, plrMinimumDepth[ y >> log2BlockSize ][ x >> log2BlockSize ] )

* + If PatchProjectionFlag[ pIdx ] is equal to 1 then,

deltaDepth = Min( deltaDepth,  
 – (plrMinimumDepth[ y >> log2BlockSize ] [ x>>log2BlockSize ]) )

* If deltaDepth is different from 0, then
  + Add a point to reconstructed geometry point cloud with same tangent and bi-tangent coordinates as the current point, and normal coordinate equal to normal coordinate of the current point plus deltaDepth.
    - The variable plrDepth is set equal to GFrame[ mapIdx ][ y ][ x ] + deltaDepth
    - The variable orgDepth is set equal to GFrame[ mapIdx ][ y ][ x ]
* Clause 9.4.8 is invoked with ( x, y ), pIdx, asps\_map\_count\_minus1, plrDepth, and GFrame as inputs. The output is a variable, isDuplicate.
* If isDuplicate is equal to 0, the following applies:

pos[ PatchAxisZ[ pIdx ] ] = plrDepth  
for( k = 0; k < 3; k++ )  
 recPcGeo[ PointCnt ][ k ] = pos[ k ]  
AttrPresent[ PointCnt ] = 0  
PointCnt ++

* + If plrFillingFlag[ y >> log2BlockSize ][ x >> log2BlockSize ] is equal to 1, then the points between orgDepth and plrDepth are filled as follows:

maxDepth = Max( orgDepth, plrDepth )  
minDepth = Min( orgDepth, plrDepth )

* For s = 1..( maxDepth – minDepth – 1), clause 9.4.8 is invoked with ( x, y ), pIdx, asps\_map\_count\_minus1, (minDepth + s ), and GFrame as inputs. The output is a variable, isDuplicate.
* If isDuplicate is equal to 0, the following applies:

pos[ PatchAxisZ[ pIdx ] ] = minDepth + s  
 for( k = 0; k < 3; k++ )  
 recPcGeo[ PointCnt ][ k ] = pos[ k ]  
 AttrPresent[ PointCnt ] = 0  
 PointCnt ++

### Reconstruction of points for non-RAW intra and inter coded patches

The inputs to this process are:

* PointCnt, the current number of points in the reconstructed point cloud frame,
* p, the patch index,
* ( u, v ), the coordinates within the patch,
* ( x, y ), the coordinates of the point on the atlas with index AtlasIdx,
* GFrame[ mapIdx ][ y ][ x ], where y = 0..asps\_frame\_height – 1, x = 0.. asps\_frame\_width – 1,
* OFrame[ y ][ x ], where y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1,
* AFrame[ aIdx ][ mapIdx ][ cIdx ][ y ][ x ], where aIdx = 0 .. ai\_attribute\_count – 1, cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ], y = 0 .. asps\_frame\_height – 1, and x = 0 .. asps\_frame\_width – 1,
* PatchProjectionFlag[ p ], the plane on which the current patch with patch index p is projected,
* PatchPackingBlockSize, the size of the block used for the horizontal and vertical placement of the patches within the atlas.

The outputs of this process are:

* [Ed. (RJ): To be completed.]

For mapIdx = 0 .. asps\_map\_count\_minus1, the following applies:

– The 3D coordinates corresponding to the point with patch coordinates ( u, v ) are derived by invoking clause 9.4.6 with patch coordinates ( u, v ), patch index pIdx, and GFrame[ mapIdx ][ y ][ x ] as inputs. The output is a three-dimensional vector, pos3D, containing the ( x, y, z ) coordinates of the point.

* + If mapIdx is equal to zero, a variable, isDuplicate, is set to 0.
* Otherwise (mapIdx is greater than 0), clause 9.4.8 is invoked with atlas coordinates ( x, y ), map index mapIdx, and GFrame as inputs. The output is the variable, isDuplicate.
* If isDuplicate is equal to 0, the following applies:
* The vector, pos3D, is copied over to a new entry in RecPcGeo.

for( n=0; n < 3; n++ )  
 RecPcGeo[ PointCnt ][ n ] = pos3D[ n ]

* For aIdx = 0 .. ai\_attribute\_count[ AtlasIdx ]− 1, attributes are assigned to the newly added point as follows:

numComps = ai\_attribute\_dimension\_minus1[ AtlasIdx ][ aIdx ] + 1  
for( cIdx = 0; cIdx  <  numComps; cIdx+ + )  
 RecPcAttr[ PointCnt ][ aIdx ][ cIdx ] = AFrame[ aIdx ][ mapIdx ][ cIdx ][ y ][ x ]

* AttrPresent[ PointCnt ] = 1
* PointCnt ++.

### Reconstruction of points when asps\_eom\_patch\_enabled\_flag is 1

Inputs to this process are:

* GFrame[ mapIdx ][ y ][ x ], where mapIdx = 0..asps\_map\_count\_minus1, y = 0..asps\_frame\_height – 1, and x = 0.. asps\_frame\_width – 1,
* OFrame[ y ][ x ], where y = 0..asps\_frame\_height – 1 and x = 0..asps\_frame\_width – 1,
* AFrame[ aIdx ][ mapIdx ][ cIdx ][ y ][ x ], where mapIdx = 0..asps\_map\_count\_minus1, y = 0 .. asps\_frame\_height – 1, x = 0 .. asps\_frame\_width – 1, aIdx = 0 .. ai\_attribute\_count – 1, and cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ].
* pIdx, the current patch index,
* PatchProjectionFlag[ pIdx ], the plane on which the current patch with patch index pIdx is projected, and
* PatchPackingBlockSize, the size of the block used for the horizontal and vertical placement of the patches within the atlas

Outputs to this process are:

* updated PointCnt,
* updated RecPcGeo[ n ][ k ], a container holding a list of coordinates of the points in the reconstructed point cloud frame, where n = 0 .. PointCnt – 1 and k = 0 .. 2, and
* updated RecPcAttr[ n ][ aIdx ][ cIdx ], a container holding a list of attributes corresponding to the points in the reconstructed point cloud frame, where n = 0 .. PointCnt – 1, aIdx = 0 .. ai\_attribute\_count – 1, and cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ].

The vector EomPointOffset of size 2afps\_eom\_number\_of\_patch\_bit\_count\_minus1 + 1, containing the relative address of each first EOM point associated to each geometry patch is derived as follows:

EomPointOffset[ epdu\_patch\_count\_minus[ pIdx ][ 0 ] ] = 0

for( i = 1; i < epdu\_patch\_count\_minus1[ pIdx ]+1; i++ )

EomPointOffset[ epdu\_associated\_patch\_idx[ pIdx ][ i ] ] =  
 EomPointOffset[ epdu\_associated\_patch\_idx[ pIdx ][ i − 1 ] ] + epdu\_points[ pIdx ][ i − 1]

[Ed. (PA): Associated patch interpretation is currently under discussion wih Jungsun. The principal question is whether the epdu\_\* syntax elements should be written with regular patch (having EOM points geometry in their occupancy map) perspective or with EOM (attribute) patch perspective. It does have an impact on associated patch semantics and derived equations. Multiple EOM (ATTR) patches should also been addressed but AFPS variables are related to all GEO/OCM patches...]

The variable eomPointCnt, indicating the number of processed EOM points in the current EOM geometry patch contained in the occupancy map, is initialized to 0.

For u = 0..Patch2dSizeX[ pIdx ] − 1, v = 0..Patch2dSizeY[ pIdx ] – 1, the following applies:

[Ed. (RJ): Need to check that BlockToPatch is p as well.][Ed. (PA): Is this comment still relevant?]

* The conversion of patch coordinates to atlas coordinates as specified in clause 9.4.5 is invoked with patch coordinates (u, v), and patch index pIdx, as inputs and atlas coordinates (x, y) as output.

– If OFrame[ y ][ x ] is not equal to 0, the following ordered steps apply:

1. A variable D0 is set equal to GFrame[ 0 ][ y ][ x ].
2. A variable D1 is derived as follows:

* If asps\_map\_count\_minus1 is equal to 0, a variable D1 is set equal to D0 + asps\_eom\_fix\_bit\_count\_minus1 + 1.
* Otherwise ( asps\_map\_count\_minus1 is greater than 0 ) D1 is set equal to GFrame[ 1 ][ y ][ x ].

1. If ( D1 – D0 ) is greater than 1, the following applies:

* A variable eomCode is derived as follows:

eomCode = ( (1 << (D1 – D0) ) – 1 ) − OFrame[ y ][ x ]

* For b = 0 .. (D1 – D0 – 2 ), if ( ( ( eomCode >> b ) & 0x01 ) = = 1 ), the following applies:
  + Clause 9.4.6 is invoked with the patch coordinates ( u, v ) , the patch index pIdx, and depth ( D0 + b + 1 ) and the output is an array pos3D of dimension 3.
  + The vector, pos3D, is copied over to a new entry in RecPcGeo.

for( n=0; n < 3; n++ )

RecPcGeo[ PointCnt ][ n ] = pos3D[ n ]

* + For aIdx = 0 .. ai\_attribute\_count[ atlasIdx ]− 1, attributes are assigned to the newly added point as follows:

A variable pIdxOcm is set equal to the index of the EOM points geometry patch contained in the occupancy map in raster scan order and corresponding to OFrame[ y ][ x ]

eomPos = EomPointOffset[ pIdxOcm ] + eomPointCnt

inBlockPos = eomPos % ( PatchPackingBlockSize \* PatchPackingBlockSize )

nbBlock = eomPos / ( PatchPackingBlockSize \* PatchPackingBlockSize )

xBlock = nbBlock % ( EomPatch2dSizeX [ pIdx ]/ PatchPackingBlockSize )

yBlock = nbBlock / ( EomPatch2dSizeY [ pIdx ]/ PatchPackingBlockSize )

xx = xBlock \* PatchPackingBlockSize + inBlockPos % PatchPackingBlockSize

+ EomPatch2dPosX[ pIdx ]

yy = yBlock \* PatchPackingBlockSize + inBlockPos / PatchPackingBlockSize

+ EomPatch2dPosY[ pIdx ]

attrDim = ai\_attribute\_dimension\_minus1[ atlasIdx ][ aIdx ] + 1

for( cIdx = 0; cIdx  <  attrDim; cIdx+ + )

RecPcAttr[ PointCnt ][ aIdx ][ cIdx ] = AFrame[ aIdx ][ 0 ][ cIdx ][ yy ][ xx ]

* + eomPointCnt ++
  + PointCnt ++

### Conversion from patch coordinates to atlas coordinates

Inputs to this process are:

* Patch coordinates (u, v)
* p, the patch index

Outputs of this process are the atlas coordinates (x, y).

The variable orientationIdx is set to PatchOrientationIndex[ p ].

The atlas coordinates (x, y) are obtained as follows:

where the outputs of functions Rotation(x) and Offset(x) are matrices as specified in Table 9‑1.

Table 9‑1 Flexible Patch Orientation

|  |  |  |  |
| --- | --- | --- | --- |
| **x** | **Identifier** | **Rotation( x )** | **Offset( x )** |
| 0 | FPO\_NULL |  |  |
| 1 | FPO\_SWAP |  |  |
| 2 | FPO\_ROT90 |  |  |
| 3 | FPO\_ROT180 |  |  |
| 4 | FPO\_ROT270 |  |  |
| 5 | FPO\_MIRROR |  |  |
| 6 | FPO\_MROT90 |  |  |
| 7 | FPO\_MROT180 |  |  |

### Reconstruct 3D point position

The inputs to this process are:

* patch coordinates (u, v)
* patch index pIdx and,
* depth value depthValue.

The output of this process is a vector, pos3D, of dimension 3.

The following steps apply:

* pos3D[ PatchAxisX[ pIdx ] ] = Patch3dPosX[ pIdx ] + u \* PatchLoDScaleX[ pIdx ]
* pos3D[ PatchAxisY[ pIdx ] ] = Patch3dPosY[ pIdx ] + v \* PatchLoDScaleY[ pIdx ]
* pos3D[ PatchAxisZ[ pIdx ] ] = Max(0, Patch3dPosMinZ[ pIdx ] +

(1 – 2\* PatchProjectionMode[ pIdx ]) \* depthValue )

### Projection planes at 45 degrees

Inputs to this process are:

* patch index, p
* three-dimensional vector point3D

The output of this process is the updated vector point3D.

The 3D coordinates of the vector point3D are adjusted based on the value of the variable Patch45degreeProjectionRotationAxis[ p ] as follows:

* The variable rotBias is initialized as follows:

rotBias = ( 1 <<  (gi\_geometry\_3d\_coordinates\_bitdepth\_minus1 +1 ) – 1 ).

* The variables rotAxis1 and rotAxis2 are derived as follows.

rotAxis1 = ( 4 − Patch45degreeProjectionRotationAxis[ p ] ) % 3  
 rotAxis2 = ( 3 − Patch45degreeProjectionRotationAxis[ p ] ).

* The 3D coordinates of the vector point3D are modified as follows:

temp1 = point3D[ rotAxis1 ]   
 temp2 = point3D[ rotAxis2 ]   
 point3D[ rotAxis1 ] = (temp1 – temp2 + rotBias ) / 2  
 point3D[ rotAxis2 ] = (temp1 + temp2 – rotBias ) / 2

### Duplicate point check

Inputs to this process are:

* atlas coordinates (xC, yC)
* curIdx, a map index
* GFrame[ mapIdx ][ y ][ x ], decoded geometry frames at nominal resolution, where mapIdx = 0 .. asps\_map\_count\_minus1, y = 0 .. asps\_frame\_height – 1, and x = 0 .. asps\_frame\_width – 1.

Outputs of this process is a variable, isDuplicate, indicating whether the point corresponding to curIdx is a duplicate of a point from the set of points with map indices in the range of 0 .. curIdx – 1.

The variable isDuplicate is assigned a value of 0.

If asps\_remove\_duplicate\_point\_enabled\_flag is equal to 1, the following applies:

for( i = 0; ( i <  curIdx && !isDuplicate ) ; i++ )  
 if( GFrame[ i ][ yC ][ xC ] = = GFrame[ curIdx ][ yC ][ xC ] )  
 isDuplicate = 1

## Reconstruction of point cloud from RAW coded patches

[Ed. (PA): Equation numbering should be added to Sections 9.4 and 9.5]

[Ed. (PA): auxiliary\_video not considered yet.We may need to define a gRawFrame when sepearate\_video is enabled. There are no such separate variables for Attributes and Geo right now. We could just define a variable change at the beginning of this clause and the process would remain the same for separate or non separate video. Something like:

when vps\_auxiliary\_video\_present\_flag[ j ] is 0

geoFrame][ y ][ x]  = gFrame[y][x],

otherwise (when vps\_auxiliary\_video\_present\_flag[ j ] is 1)

geoFrame][ y ][ x]  = gRawFrame[ y ][ x ]

and then replace gFrame by geoFrame in the current clause.]

[Ed. (PA): Add a figure to show RAW points layout.]

Inputs to this process are:

* a variable pointCnt specifying the number of points in the reconstructed point cloud frame.
* an array gFrame[ y ][ x ], specifying geometry information, where y is in the range of 0 to asps\_frame\_height – 1, inclusive, and, x is in the range of 0 to asps\_frame\_width – 1, inclusive.
* AFrame[ aIdx ][ 0 ][ cIdx ][ y ][ x ], where y = 0 .. asps\_frame\_height – 1, x = 0 .. asps\_frame\_width – 1, aIdx = 0 .. ai\_attribute\_count – 1, and cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ].
* when ai\_attribute\_count[ AtlasIdx ] is greater than 0, an array aFrame[ aIdx ][ 0 ][ cIdx ][ y ][ x ], specifying attribute information, where aIdx is in the range of  0 to ai\_attribute\_count – 1, inclusive, cIdx is in the range of  0 to ai\_attribute\_dimension\_minus1[ aIdx ], inclusive, y is in the range of 0 to asps\_frame\_height – 1, inclusive, and x is in the range of 0 to asps\_frame\_width – 1, inclusive,
* an array recPcGeo[ n ][ k ], specifying a list of coordinates of the points in the reconstructed point cloud frame, where n is in the range of 0 to pointCnt – 1, inclusive, and k is in the range of 0 to 2, inclusive,
* when ai\_attribute\_count[ AtlasIdx ] is greater than 0, an array recPcAttr[ n ][ attrIdx ][ compIdx ], specifying a list of attributes associated with the points in the reconstructed point cloud frame, where n is in the range of 0 to pointCnt – 1, inclusive, attrIdx is in the range of 0 to ai\_attribute\_count – 1, inclusive, and compIdx is in the range of 0 to ai\_attribute\_dimension\_minus1[ attrIdx ], inclusive, and

– a variable pIdx, specifying the index of the patch.

Outputs of this process are the updated variable pointCnt and updated arrays recPcGeo and recPcAttr.

updated variable pointCnt

* rp
* updated recPcAttr[ n ][ aIdx ][ cIdx ], a container holding a list of attributes corresponding to the points in the reconstructed point cloud frame, where n = 0 .. pointCnt – 1, aIdx = 0 .. . ai\_attribute\_count – 1, and cIdx = 0 .. ai\_attribute\_dimension\_minus1[ aIdx ].

[Ed. (AMT): Need to revise this text based on the vps\_raw\_patch\_enabled\_flag ]

The geometry of the RAW coded points is retrieved from the RAW coded patch in the geometry frame and added to RecPcGeo, as specified below:

The variables numRawPoints, xRawPos, yRawPos, zRawPos are initialized as follows:

numRawPoints = 0

xRawPos = PatchRawPoints[ pIdx ]

yRawPos = 2 \* PatchRawPoints[ pIdx ]

zRawPos = 3 \* PatchRawPoints[ pIdx ]

for( x = Patch2dPosX[ pIdx ]; x < Patch2dPosX[ pIdx ] + Patch2dSizeX[ pIdx ]; x++ )

for( y = Patch2dPosY[ pIdx ]; y < Patch2dPosY[ pIdx ] + Patch2dSizeY[ pIdx ]; y++ ) {

if( numRawPoints < xRawPos )

RecPcGeo[ pointCnt ][ 0 ] = gFrame[ y ][ x ] + Patch3dPosX[ pIdx ]

else if ( numRawPoints  < yRawPos )

RecPcGeo[ pointCnt ][ 1 ] = gFrame[ y ][ x ] + Patch3dPosY[ pIdx ]

else if (numRawPoints  < zRawPos )

RecPcGeo[ pointCnt ][ 2 ] = gFrame [ y ][ x ] + Patch3dPosZ[ pIdx ]

numRawPoints ++

pointCnt ++

}

[Ed. (PA): PointCnt should be specified for ATTR or GEO. Should we have PointCntAttr, PointCntGeo?]

While the 3-tuple (x, y, z) coordinates of the RAW coded points are stored in a single patch in the first component of the geometry frame, the attribute information for a RAW coded point is stored across the different attribute dimensions in the attribute frame at the same location. The attribute information of the RAW coded points is retrieved from the different attribute dimensions of the attribute frame and added to RecPcAttr as specified below:

=

pointCntAttr = PointCnt − numRawPoints

for(j = 0; j < vps\_atlas\_count\_minus1 + 1 ; j++ )

for( x = Patch2dPosX[ pIdx ]; x < Patch2dPosX[ pIdx ] + Patch2dSizeX[ pIdx ]; x++ )

for( y = Patch2dPosY[ pIdx ]; y < Patch2dPosY[ pIdx ] + Patch2dSizeY[ pIdx ]; y++ ) {

if( numRawPoints  < PatchRawPoints[ pIdx ] ) {

for( aIdx = 0; aIdx < ai\_attribute\_count ; aIdx++ ) {

attrDim = ai\_attribute\_dimension\_minus1[ j ][ aIdx ] + 1

for( cIdx = 0; cIdx < attrDim; cIdx++ )

RecPcAttr[ PointCnt ][ aIdx ][ cIdx ] = AFrame[ aIdx ][ 0 ][ cIdx ][ y ][ x ]

}

AttrPresent[ pointCntAttr ] = 1

numRawPoints ++

pointCntAttr ++

}

}

[Ed. (PA): Instead of looping over atlas, it may be enough to say that for the current atlas with index …]

## Smoothing process

### General

This process is invoked after reconstructing the geometry and attributes for the current atlas with index atlasIdx.

The inputs to this process are:

* pointCnt, the number of reconstructed points,
* recPcGeo[ n ][ k ], n = 0..pointCnt – 1, k = 0..2, the reconstructed geometry, and
* recPCAttr[ n ][ i ][ j ], n = 0..pointCnt – 1, i = 0..ai\_attribute\_count[ atlasId ] – 1, and j = 0..ai\_attribute\_dimension\_minus1[ atlasId ][ i ], the reconstructed attributes.

The outputs of this process are:

* recPcGeoSmoothed
* recPcAttrSmoothed
* IsBoundaryPoint[ n ], n =0..PointCnt – 1, an array specifying whether the reconstructed point lies on or near patch boundary on the atlas.

The following ordered steps apply:

1. An array, IsBoundaryPoint[ n ], n =0..pointCnt – 1, is initialized to 0.
2. If a bitstream contains a geometry smoothing SEI message with a value of gs\_smoothing\_instance\_type[ k ] equal to 1, and the point cloud reconstruction system has selected to use this k-th geometry smoothing instance for performing smoothing during point cloud reconstruction, clause 9.6.2 is invoked. (specify input and outputs)
3. Otherwise, if a bitstream contains a geometry smoothing SEI message with a value of gs\_smoothing\_instance\_type[ k ] equal to 2, and the point cloud reconstruction system has selected to use this k-th geometry smoothing instance for performing smoothing during point cloud reconstruction, Clause 9.6.8 is invoked. (specify input and outputs)
4. If the point cloud reconstruction system has selected to use the k-th geometry smoothing instance for which gs\_smoothing\_instance\_type[ k ] is greater than 0, attribute transfer as specified in clause A.5.1.4 may be invoked with pointCnt, IsBoundaryPoint[ n ], n = 0..pointCnt – 1, and recPcGeoSmoothed as inputs (specify outputs). This is a soft confromance requirement. Clause A.5.1.4 may be substituted by an alternate process that produces similar visual quality. The definition of same visual quality is outside the scope of this document.
5. For each attribute j, j =0..ai\_attribute\_count[ atlasId ] – 1, if a bitstream contains an attribute smoothing SEI message with a value of as\_smoothing\_filter\_instance\_type[ j ][ i ] equal to 1, and the point cloud reconstruction system has selected to use this i-th attribute smoothing instance for performing attribute smoothing for attribute j during point cloud reconstruction, clause 9.6.3 is invoked. (specify input and outputs)

[Ed. (RJ) (ATA): Should we add a note as to how the filter instance is chosen by the point cloud reconstruction system. Maybe add a note.]

### Geometry smoothing process

This process is invoked when a bitstream contains a geometry smoothing SEI message with a value of gs\_smoothing\_instance\_type[ r ] equal to 1, and the point cloud reconstruction system has selected to use this r-th geometry smoothing instance for performing smoothing during point cloud reconstruction.

[Ed. (AMT): We may wish to move this section into an Annex instead of this being part of section 9. The same applies to attribute smoothing. Also, since we now can support multiple methods for smoothing, we may wish to add a general section and this is then followed with the different supported methods.]

Inputs to this process are:

* oFrame[ y ][ x ], y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1, an occupancy map frame,
* gridSize, indicating the geometry smoothing grid size,
* geoSmoothingThreshold, indicating the geometry smoothing threshold,
* recPcGeo[ y ][ x ], y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1, an array holding the list of points in the reconstructed point cloud geometry frame, and
* pointCnt, the number of points in the current reconstructed point cloud frame

[Ed. (RJ): It may be better to use the term geometry grid smoothing to refer to this method, eventhough there is no other geometry smoothing method in the specification. As discussed the other method should be referred to as occupancy map smoothing. If we decide to use the term geometry grid smoothing, some of the syntax element names should be changed as well.]

[Ed. (RJ): As discussed in the F2F AHG meeting, EOM points should perhaps also be excluded from smoothing.]

Outputs of this process are:

* recPcGeoSmoothed[ n ], n = 0..pointCnt – 1, an array holding the list of points in the reconstructed point cloud geometry frame after the geometry smoothing process, and
* isBoundaryPoint[ n ], n = 0..pointCnt – 1, an array specifying whether the geometry smoothing process affected the position of the reconstructed point with index n.

The variable isBoundaryPoint[ n ] may take values of 0, 1, or 2. A boundaryPoint value of 0 indicates that a point is not near a patch boundary. A value of 1 indicates that the reconstructed point is on or near a patch boundary. A value of 2 indicates that the position of the reconstructed point is changed by the geometry smoothing process.

The following applies:

* An array isBoundaryPoint[ n ], n =0..pointCnt – 1, is initialized to 0.
* An array recPcGeoSmoothed[ n ], n = 0..pointCnt – 1, using the array recPcGeo[ n ], n = 0..pointCnt – 1.
* For each reconstructed point with index n, n =0..pointCnt – 1, the following applies:
* If oFrame[ y ][ x ] is not equal to 0 and isBoundaryPoint[ n ] is equal to 0, clause 9.6.5 is invoked with oFrame, PointToPixel[ n ][ 0 ], and PointToPixel[ n ][ 1 ] as inputs. The output is assigned to isBoundaryPoint[ n ].

[Ed. (RJ): Currently, this may classify EOM points as boundary points and apply smoothing to them. This should be corrected. Would need a contribution.]

[Ed. (RJ): Set PointToPixel value in clauses 9.3 and 9.4. A value of −1 may be used to signify RAW or EOM coded point.]

* Clause 9.6.4.1 is invoked with inputs gridSize, pointCnt, isBoundaryPoint[ n ], recPcGeo[ n ][ k ], and PointToPatch[ n ] for n = 0..pointCnt – 1, k = 0..2, and the outputs are numCells1D, numBoundaryCells, cellIdxLut[ x ][ y ][ z ], for x, y, and z in the range of 0 to numCells1D – 1, geoCenterGrid[ m ][ k ], cellCnt[ m ], cellDoSmoothing[ m ] for m = 0..numBoundaryCells – 1, k = 0..2.

[Ed. (RJ): Set PointToPatch value in clauses 9.3 and 9.4.

* For each reconstructed point with index n, n =0..pointCnt – 1
* The following applies:

x = recPcGeo[ n ][ 0 ]  
y = recPcGeo[ n ][ 1 ]  
z = recPcGeo[ n ][ 2 ]

halfGridSize = Max( gridSize / 2, 1 );  
maxCoord = gridSize \* numCells1D;  
if( x < halfGridSize || y < halfGridSize || z < halfGridSize || maxCoord <= x + halfGridSize ||  
 maxCoord <= y + halfGridSize || maxCoord <= z + halfGridSize )  
 continue

* If isBoundaryPoint[ n ] is true, clause 9.6.6 is invoked with numCells1D, numBoundaryCells, recPcGeo[ n ][ k ], k =0..2, cellIdxLut, cellCnt, and cellDoSmoothing as inputs and the outputs are otherClusterPtCnt, idxArr[ dx ][ dy ][ dz ], dx, dy, and dz in the range of 0 to 1, inclusive, and cellAnchor[ k ], k = 0..2.
* When otherClusterPointCnt is equal to 1, the following applies:
* A 3-D array centroid[ x ][ y ][ z ], for x, y, z in the range of 0 to 1, inclusive, is calculated as follows:

for( dx = 0; dx < 2; dx++ )  
 for( dy = 0; dy < 2; dy++ )  
 for( dz = 0; dz < 2; dz++ ) {  
 tmpIdx = idxArr[ dx ][ dy ][ dz ]  
 if( ( cellCnt[ tmpIdx ] > 0 ) && ( idxArr[ dx ][ dy ][ dz ] != −1 ) )  
 for( k = 0; k < 3; k++ )  
 centroid[ dx ][ dy ][ dz ][ k ] = geoCenterGrid[ tmpIdx ][ k ]  
 else  
 for( k = 0; k < 3; k++ )  
 centroid[ dx ][ dy ][ dz ][ k ] = recPcGeo[ n ][ k ]  
 }

* Clause 9.6.7 is invoked with gridSize, numBoundaryCells, recPcGeo[ n ][ k ], k = 0..2, centroid[ x ][ y ][ z ][ k ], for x, y, and z in the range of 0 to 1, inclusive, and k =0..2, cellAnchor[ k ], k = 0..2, idxArr[ dx ][ dy ][ dz ], dx, dy, and dz in the range of 0 to 1, and cellCnt[ n ], n = 0..numBoundaryCells – 1 as inputs and the output is pos3dFilt[ k ], k =0..2, a filtered version of the current reconstructed point.
* The following applies:

sqrNorm = 0  
for( k = 0; k < 3; k++ ) {  
 diff = recPcGeo[ n ][ k ] \* cnt − centroid[ k ]  
 sqrNorm += diff \* diff  
}  
dist2 = sqrNorm ÷ cnt + 0.5  
if( dist2 >= Max( geoSmoothingThreshold, cnt ) \* 2 ) {  
 for( k = 0; k < 3; k++ )  
 recPcGeoSmoothed[ n ][ k ] = Rnd( centroid[ k ] ÷ cnt )  
 isBoundaryPoint[ n ] = 2  
}

### Attribute smoothing process

This process is invoked for the current atlas, with index AtlasIdx when attribute type is texture, attribute dimension is equal to 1 or 3, a bitstream contains an attribute smoothing SEI message with a value of as\_smoothing\_method\_type[ k ][ i ] equal to 1, and the point cloud reconstruction system has selected to use filtering method i for performing smoothing during point cloud reconstruction of the attribute with index k.

[Ed. (AMT): We may wish to move this section into an Annex instead of this being part of section 9. Also, since we now can support multiple methods for smoothing, we may wish to add a general section and this is then followed with the different supported methods.]

Inputs to this process are:

* an occupancy map corresponding to the current point cloud frame at nominal resolution, OFrame,
* an attribute index, aIdx,
* the current smoothing filter method mIdx,
* number of components, numComps, corresponding to the attribute index aIdx,
* an attribute smoothing control parameter set, AttributeSmoothingGridSize[ aIdx ][ mIdx ], as\_smoothing\_threshold[ aIdx ][ mIdx ], as\_smoothing\_threshold\_variation[ aIdx ][ mIdx ], and as\_smoothing\_threshold\_difference[ aIdx ][ mIdx ]
* an array containing reconstructed attribute values for attribute index aIdx, RecPcAttrSmIn[ aIdx ][ i ][ j ], , ,
* an array containing reconstructed (possibly smoothed) reconstucted positions, RecPcGeomSmOut[ i ][ j ], , and
* patch index information corresponding to each point in RecPcGeomSmOut

Output of this process is:

* an array containing reconstructed smoothed attribute values for attribute index aIdx, RecPcAttrSmOut[ aIdx ][ i ][ j ], , .

The variable GridSize is set to AttributeSmoothingGridSize[ aIdx ][ mIdx ].

Number of cells, numCells, and the arrays of attrCenterGrid[ i ] and cellDoSmoothing[ x ][ y ][ z ], for i = 0 to numCells -1, inclusive, are derived as described in 9.6.3.2:

for i = 0 to ( PointCnt – 1 ), inclusive, the following applies:

* otherClusterPtCnt is set equal to 0.

To determine the points at the patch boundary, clause 9.6.4 is invoked to produce an array recPCBoundary[ idx ], , which identifies whether each point from RecPcGeomSmOut is a boundary point.

[Ed (PA): I think the part just above should be adjusted to 9.6.5. There is no more recPCBoundary. Same for PBF in 9.6.8]

for i = 0 to ( PointCnt – 1 ), inclusive, the following applies:

* If recPCBoundary[ i ] is equal to 1 the following steps are performed:
* Variables pointGeom[ j ], j in the range of 0 to 2, inclusive, and pointAttr[ k ], k in the range of 0 to ( numComps – 1), inclusive, are defined as follows:

for( j = 0; j < 3 ; j++ )  
 pointGeom[ j ] = RecPcGeomSmOut[ i ][ j ]  
 for( k = 0; k < numComps ; k++ )  
 pointAttr[ k ] = RecPcAttrSmIn[ aIdx ][ i ][ k ]

sub clause 9.6.5 is applied to derive:

* a 2x2x2 grid neighbourhood corresponding to the current position, pointGeom[ i ]
* the top-left corner of 2x2x2 grid, s[  i ],
* the2x2x2 grid position associated with the current position, t[ i ], i = 0 to 2, inclusive
* a boolean value otherClusterPtCn

If otherClusterPtCnt is equal to 1, the following applies:

* The variation of attribute in the 2×2×2 cell neighbourhood of the current cell is determined as follows:

The isOriginalCell array for the 2×2×2 cell neighbourhood is determined as follows:

for( dx = 0; dx < 2; dx++ ) {  
 for( dy = 0; dy < 2; dy++ ) {   
 for( dz = 0; dz < 2; dz++ ) {  
 xIdx = s[ 0 ] + dx  
 yIdx = s[ 1 ] + dy  
 zIdx = s[ 2 ] + dz  
 if( ( xIdx == t[0] ) && ( yIdx == t[1] ) && ( zIdx == t[2] ) ) {   
 isOriginalCell[ dx ][ dy ][ dz ] = 1   
 for( k=0; k < numComps; k++ ) {    
 currAttr[ k ] = pointAttr[ k ]   
 }  
 if( abs( meanLuma[ xIdx ][ yIdx ][ zIdx ] − medianLuma[ xIdx ][ yIdx ][ zIdx ] )  
 <= as\_smoothing\_threshold\_variation[ aIdx ][ mIdx ] ) {   
 lumaOrig = centroidLuma[ xIdx ][ yIdx ][ zIdx ]   
 for( k=0; k < numComps; k++ ) {   
 attrCentroid[ dx ][ dy ][ dx ][ k ] =   
 attrCenterGrid[ xIdx ][ yIdx ][ zIdx ][ k ]  
 }  
 }   
 else {   
 lumaOrig = currAttr[ 0 ]   
 for( k=0; k < numComps; k++ ) {   
 attrCentroid[ dx ][ dy ][ dx ][ k ] = currAttr[ k ]  
 }  
 }  
 }  
 else {  
 isOriginalCell[ dx ][ dy ][ dz ] = 0  
 }  
 }  
 }  
 }

* The variation in attribute values within a cell and difference between luma centroid values of the original cell and neighboring cells is determined as follows:

for( dx = 0; dx < 2; dx++ ) {   
 for( dy = 0; dy < 2; dy++ ) {  
 for( dz = 0; dz < 2; dz++ ) {  
 xIdx = s[ 0 ] + dx  
 yIdx = s[ 1 ] + dy  
 zIdx = s[ 2 ] + dz  
 if( isOriginalCell[ dx ][ dy ][ dz ] == 0 ) {  
 if( cellCnt[ xIdx ][ yIdx ][ zIdx ] > 0 ) {  
 for( k=0; k < numComps; k++ ) {  
 attrCentroid[ dx ][ dy ][ dx ][ k ] =   
 attrCenterGrid[ xIdx ][ yIdx ][ zIdx ][ k ]  
 }  
 lumaN = centroidLuma[ xIdx ][ yIdx ][ zIdx ]  
 diff = abs( lumaOrig – lumaN )  
 var = meanLuma[ xIdx ][ yIdx ][ zIdx ]   
 − medianLuma[ xIdx ][ yIdx ][ zIdx ]  
 if( ( diff > as\_smoothing\_threshold\_difference[ aIdx ][ mIdx ]  ) ||  
 ( abs( var ) > as\_smoothing\_threshold\_variation[ aIdx ][ mIdx ] ) ) {  
 for( k = 0; k < numComps; k++ ) {  
 attrCentroid[ dx ][ dy ][ dx ][ k ] = currAttr[ k ]  
 }  
 }  
 }   
 else {  
 for( k=0; k < numComps; k++ )  
 attrCentroid[ dx ][ dy ][ dx ][ k ] = currAttr[ k ]  
 }  
 }  
 }  
 }  
 }

* An 8-tap trilinear filter as described in clause 9.6.6 is applied to the attribute centroids, attrCentroid[ x ][ y ][ z ][ k ], where , , in the 2×2×2 neighborhood as described in clause 9.6.5 is applied:
* If the following distortion criteria are satisfied, the output of the tri-linear filter will be added to recPCaSmOut.
* The variable distToCentroid2 defined as abs(Ycell – Ycur) \* 10 is greater than as\_smoothing\_threshold[ aIdx ][ mIdx ] , where Ycell and Ycur are the luma values of the centroid of points in the current cell (attrCentroid) and the current point (pointAttr), respectively.

### Identification of the Center Grid

#### Identification of the geometry center grid

Inputs to this process are:

* gridSize, indicating the geometry smoothing grid size,
* pointCnt, the number of points in the current reconstructed point cloud frame,
* isBoundaryPoint[ n ], n = 0..pointCnt – 1, indicating whether the n-th reconstructed point lies on or near a patch boundary,
* recPcGeo[ i ][ k ], i = 0..pointCnt – 1, k = 0..2, an array containing the reconstructed point positions, and
* pointToPatch[ n ], n = 0..pointCnt – 1, an array containing the index of the patch that the n-th reconstructed point, recPcGeo[ n ], belongs to.

Output of this process are:

* numCells1D, the number of cells in x, y, or z direction,
* numBoundaryCells, the number of cells containing a boundary point or neighboring such a cell,
* cellIdxLut[ x ][ y ][ z ], for x, y, and z in the range of 0 to numCells1D − 1, inclusive, an array mapping a 3-D cell index to a 1-D index in the range of 0..numBoundaryCells – 1,
* centerGrid[ i ][ k ], i = 0..numBoundaryCells – 1 and k = 0..2, an array containing reconstructed geometry center grid positions,
* cellCnt[ n ], n = 0..numBoundaryCells – 1, an array containing number of reconstructed points in cells, and
* cellDoSmoothing[ n ], n = 0..numBoundaryCells – 1, a boolean array indicating whether smoothing should be performed for a cell.

The following applies:

* The three-dimensional geometry coordinate space is divided into cells of size gridSize × gridSize × gridSize. The variable numCells1D is derived as follows:

for( k = 0; k < 3; k++ ) {  
 maxG[ k ] = recPcGeo[ 0 ][ k ]  
}

for( i = 1; i < pointCnt ; i++ )  
 for( k= 0; k < 3 ; k++ ) {  
 maxG[ k ] = Max(maxG[ k ], recPcGeo[ i ][ k ] )   
numCells1D = ( Max( maxG[ 0 ], Max( maxG[ 1 ], maxG[ 2 ] ) ) + gridSize − 1 ) / gridSize

* Clause 9.6.4.3 is invoked with numCells1D, gridSize, pointCnt, isBoundaryPoint[ n ], n = 0..pointCnt – 1, and recPcGeo[ i ][ k ], i = 0..pointCnt – 1, k = 0..2, as inputs and the outputs are an array cellIdxLut[ x ][ y ][ z ], where x, y, and z are in the range of 0 to numCells1D – 1, inclusive, and numBoundaryCells.
* The arrays cellDoSmoothing[ n ], cellCnt[ n ], cellPatchIdx[ n ], and centerGrid[ n ][k] are initialized to 0 for all x, y, and z in the range of 0 to (numBoundaryCells − 1), inclusive, and k in the range of 0 to 2, inclusive.
* For n = 0..pointCnt – 1, the following applies:
  + Variables xIdx, yIdx, zIdx, and cellIndex are derived as follows:

xIdx = recPcGeo[ n ] [ 0 ] / gridSize   
yIdx = recPcGeo[ n ][ 1 ] / gridSize   
zIdx = recPcGeo[ n ][ 2 ] / gridSize   
cellIndex = cellIdxLut[xIdx + yIdx \* numCells1D + zIdx \* numCells1D \* numCells1D ]

* + If cellIndex is not equal to −1, the following applies:
  + If cellCnt[ cellIndex ] is equal to 0, cellPatchIdx[ cellIndex ] is set to the patch index of the patch that the current reconstruction point, recPcGeo[ n ], belongs to.
  + Otherwise, if cellDoSmoothing[ cellIndex ] is 0 and cellPatchIdx[ cellIndex ] is not equal to the index of the patch that the current point, recPcGeom[ n ], belongs to, cellDoSmoothing[ cellIndex ] is set to 1.
  + The following applies:

for( k = 0; k < 3 ; k++ )  
 centerGrid[ cellIndex ][ k ] += recPcGeom[ n ][ k ]  
cellCnt[ cellIndex ]++

* For n = 0..numBoundaryCells – 1 the following applies:

if( cellCnt[ n ] > 0 )  
 for( k = 0; k < 3 ; k++ )  
 centerGrid[ n ][ k] = centerGrid[ n ][ k ] ÷ cellCnt[ n ]

#### Identification of the Attribute Center Grid

Inputs to this process are:

* an array containing reconstructed positions, RecPcGeom[ i ], with i in the range of 0 to PointCnt – 1, inclusive,
* aIdx, the attribute index,
* an array containing reconstructed attributes, RecPcAttr[aIdx][ i ], with i in the range of 0 to PointCnt – 1, inclusive, and
* numComps, indicating the number of attribute components.

Outputs of this process are:

* an array containing reconstructed center grid attribute values, attrCenterGrid[ i ][ k ], with i in the range of 0 to numCells – 1, inclusive, and k in the range of 0 to numComps – 1, inclusive,
* an array containing reconstructed center grid mean luma values, meanLuma[ k ], with k in the range of 0 to numComps – 1, inclusive, and
* an array containing reconstructed center grid luma median values, attrMean[ i ][ k ] with i in the range of 0 to numCells – 1, inclusive, and k in the range of 0 to numComps – 1, inclusive.

[Ed. (RJ): Why is the variable storing the medians named attrMean?]

The following applies:

* The elements of the arrays attrCenterGrid[ x ][ y ][ z ][ m ] and meanLuma[ x ][ y ][ z ] are initialized to 0 for all x, y, and z in the range of 0 to numCells − 1, inclusive and m in the range of 0 to numComps – 1, inclusive.
* Clause 9.6.3.1 is applied to obtain the arrays cellDoSmoothing[ x ][ y ][ z ], cellCnt[ x ][ y ][ z ], and cellPatchIdx[ x ][ y ][ z ] for all x, y, and z in the range of 0 to numCells − 1, inclusive. [Ed. (RJ): Should we specify ranges for x, y, and z, separately?]
* For i = 0 .. PointCnt – 1, the following applies:
* The variables xIdx, yIdx, and zIdx are derived as follows:

xIdx = (RecPcGeom[ i ][ 0 ] / GridSize)  
 yIdx = (RecPcGeom[ i ][ 1 ] / GridSize)  
 zIdx = (RecPcGeom[ i ][ 2 ] / GridSize)

* If cellDoSmoothing[ xIdx ][ yIdx ][ zIdx ] is equal to 1, the following applies:

for( k = 0; k < numComps ; k++ )  
 attrCenterGrid[ xIdx ][ yIdx ][ zIdx ][ k ] += RecPcAttr[aIdx][i][ k ];

If cellCnt[ x ][ y ][ z ] is greater than 0 for x, y, and z in the range of 0 to numCells −1, inclusive, [Ed. (RJ): Should we specify ranges for x, y, and z, separately?] the mean and median luma values of attribute with index aIdx, for points belonging to that cell are calculated and assigned to arrays meanLuma[ x ][ y ][ z ] and medianLuma[ x ][ y ][ z ], respectively. Also, the luma value of attrCenterGrid is assigned to centroidLuma[ x ][ y ][ z ].

[Ed. (RJ): It should be clarified how the median is calculated for even number of samples.]

[AJT: need to add median calculation. Also what is the difference between meanLuma and centroidLuma?

meanLuma[ xIdx ][ yIdx ][ zIdx ] += RecPcAttr[ aIdx ][ i ][ 0 ]

Derive attribute center grids.

for( k = 0; k < numComps ; k++ )  
 attrCenterGrid[ xIdx ][ yIdx ][zIdx ][ k] =   
 attrCenterGrid[ xIdx ][ yIdx ][ zIdx ][ k ] ÷ cellCnt[ xIdx ][ yIdx ][ zIdx ]  
meanLuma[ xIdx ][ yIdx ][zIdx ] = meanLuma[ xIdx ][ yIdx ][ zIdx ] ÷ cellCnt[ xIdx ][ yIdx ][ zIdx ]

[Ed. (RJ): There seem to be some problems with the indentation and formatting as well as inconsistency in style (active vs passive).]

#### Generation of index lookup table

Inputs to this process are:

* numCells1D, the number of cells in x, y or z direction,
* gridSize, the size of the cell in x, y or z direction,
* pointCnt, the number of reconstructed points,
* isBoundaryPoint[ n ], n = 0..pointCnt – 1, indicating whether the n-th reconstructed point lies on or near a patch boundary, and
* recPcGeo[ n ][ k ], n = 0..pointCnt – 1, k = 0..2, an array containing the reconstructed point positions.

Outputs of this process are:

* cellIdxLut[ x ][ y ][ z ], for x, y, and z in the range of 0 to numCells1D − 1, inclusive, an array mapping a 3-D cell index to a 1-D index in the range of 0..numBoundaryCells – 1, and
* numBoundaryCells, the number of cells relevant for geometry smoothing or attribute smoothing.

All the cells in the numCells1D × numCells1D × numCells1D grid are not relevant to performing geometry grid smoothing or attribute smoothing. So for storing quantities specific to a grid cell, this process creates a lookup table from 3D cell index into a smaller 1D cell structure which stores only cells that are relevant to geometry grid smoothing or attribute smoothing.

An array cellIdxLut of size numCells1D × numCells1D × numCells1D is initialized to −1 and a variable numBoundaryCells is initialized to 0.

The following applies:

halfGridSize = Max( gridSize / 2, 1 );  
maxCoord = gridSize \* numCells1D;  
for( n = 0; n < pointCnt; n++ ) {  
 if( isBoundaryPoint[ n ] = = 1 ) {  
 x = recPcGeo[ n ][ 0 ]  
 y = recPcGeo[ n ][ 1 ]  
 z = recPcGeo[ n ][ 2 ]  
  
 if( x < halfGridSize || y < halfGridSize || z < halfGridSize || maxCoord <= x + halfGridSize ||  
 maxCoord <= y + halfGridSize || maxCoord <= z + halfGridSize )   
 continue  
  
 x2 = x / gridSize  
 y2 = y / gridSize  
 z2 = z / gridSize  
  
 x3 = x % gridSize  
 y3 = y % gridSize  
 z3 = z % gridSize  
  
 qx = x2 + ( ( x3 < gridSize / 2 ) ? −1 : 0  
 qy = y2 + ( ( y3 < gridSize / 2 ) ? −1 : 0  
 qz = z2 + ( ( z3 < gridSize / 2 ) ? −1 : 0  
  
 for( ix = qx; ix < qx + 2; ix++ ) {  
 for( iy = qy; iy < qy + 2; iy++ )   
 for( iz = qz; iz < qz + 2; iz++ ) {  
 cellIndex = ix + iy \* numCells1D + iz \* numCells1D \* numCells1D  
 if( cellIdxLut[ cellIndex ] = = −1 ) {  
 cellIdxLut[ cellIndex ] = numBoundaryCells;  
 numBoundaryCells++  
 }  
 }

NOTE – This process allows various arrays such as cell centroids, smoothing flags to be stored with reduced memory requirements. However this process is not normative. A conformant decoder may choose to store the full three dimensional arrays or use a different technique for reducing the memory requirements.

### Identification of Boundary Points

Inputs to this process are:

* an occupancy map frame oFrame[ i][ j ], i = 0..asps\_frame\_height – 1, j = 0..asps\_frame\_width – 1,
* x, the column index of the reconstructed point on the current atlas, and
* y, the row index of the reconstructed point on the current atlas

Output of this process is:

* boundaryPoint, a boolean variable indicating whether the reconstructed point is relevant to geometry or attribute smoothing.

This process is invoked from the geometry or the attribute smoothing process specified in clauses 9.6.2 and 9.6.3 respectively.

A variable boundaryPoint identifies if a point is located near a patch boundary.

The following applies:

boundaryPoint = 0;  
if( ( oFrame[ y ][ x − 1 ] = = 0 ) | | ( oFrame[ y ][ x + 1 ] = = 0 ) | |  
 ( oFrame[ y − 1 ][ x ] = = 0 ) | | ( oFrame[ y + 1 ][ x ] = = 0 ) | |  
 ( oFrame[ y − 1 ][ x − 1 ] = = 0 ) | | ( oFrame[ y − 1 ][ x + 1] = = 0 ) | |  
 ( oFrame[ y + 1 ][ x − 1 ] = = 0 ) | | ( oFrame[ y + 1 ][ x + 1] = = 0 ) )  
 boundaryPoint = 1

if( boundaryPoint ! = 1 ) )  
 if( ( oFrame[ y − 2 ][ x − 2 ] = = 0 ) | | ( oFrame[ y − 2 ][ x – 1 ] = = 0 ) | |  
 ( oFrame[ y − 2 ][ x ] = = 0 ) | | ( oFrame[ y − 2 ][ x + 1 ] = = 0 ) | |  
 ( oFrame[ y − 2 ][ x + 2 ] = = 0 ) | | ( oFrame[ y − 1 ][ x – 2 ] = = 0 ) | |  
 ( oFrame[ y − 1 ][ x + 2 ] = = 0 ) | | ( oFrame[ y ][ x – 2 ] = = 0 ) | |  
 ( oFrame[ y ][ x + 2 ] = = 0 ) | | ( oFrame[ y + 1 ][ x – 2 ] = = 0 ) | |  
 ( oFrame[ y + 1 ][ x + 2 ] = = 0 ) | | ( oFrame[ y + 2 ][ x – 2 ] = = 0 ) | |  
 ( oFrame[ y + 2 ][ x – 1 ] = = 0 ) | | ( oFrame[ y + 2 ][ x ] = = 0 ) | |  
 ( oFrame[ y + 2 ][ x + 1 ] = = 0 ) | | ( oFrame[ y + 2 ][ x + 2 ] = = 0 ) )  
 boundaryPoint = 1

In the above assignment process, oFrame[ i ][ j ] is assumed to be equal to 0 if i is not in the range of 0 to ( asps\_frame\_height – 1 ), inclusive, or j is not in the range of 0 to ( asps\_frame\_width – 1 ), inclusive.

### Identification of 2 × 2 ×2 cell for a reconstructed point position

Inputs to this process are:

* numCells1D, the number of cells in x, y or z direction,
* numBoundaryCells, the number of cells containing a boundary point or neighboring such a cell,
* recPos3D[ k ], k = 0..2, an array containing the position of the current reconstructed point,
* cellIdxLut[ x ][ y ][ z ], for x, y, and z in the range of 0 to numCells1D − 1, inclusive, an array mapping a 3-D cell index to a 1-D index in the range of 0..numBoundaryCells – 1,
* cellCnt[ n ], n = 0..numBoundaryCells – 1, an array containing number of reconstructed points in a cell, and
* cellDoSmoothing[ n ], n = 0..numBoundaryCells – 1, a boolean array indicating whether smoothing should be performed for a cell.

Output of this process are:

* otherClusterPtCnt, a boolean variable indicating whether tri-linear filtering should be performed for the cell to which the current reconstructed point belongs to,
* idxArr[ dx ][ dy ][ dz ], dx, dy, and dz in the range of 0 to 1, inclusive, an array storing the 1-D indices for the 2 × 2 ×2 grid of cells, and
* cellAnchor[ k ], k = 0..2, an array specifying anchor point for the 2 × 2 ×2 grid of cells.

The 2 × 2 ×2 grid of cells neighboring the reconstructed point recPos3D[ k ], k = 0..2, is identified as follows:

for( k = 0; k < 3 ; k++ ) {  
 idxCell[ k ] = ( recPos3D[ k ] / gridSize )  
 cellAnchor[ k ] = idxCell[ k ] + ( ( ( recPos3D[ k ] % gridSize ) < ( gridSize / 2 ) ) ? –1 : 0 )  
}  
  
otherClusterPtCnt = 0  
for( dx = 0; dx < 2; dx++ ) {  
 for( dy = 0; dy < 2; dy++ ) {  
 for( dz = 0; dz < 2; dz++ ) {  
 xIdx = cellAnchor[ 0 ] + dx  
 yIdx = cellAnchor[ 1 ] + dy  
 zIdx = cellAnchor[ 2 ] + dz  
 if  
 tmpIdx = cellIdxLut[ xIdx ][ yIdx ][ zIdx ]  
 idxArr[ dx ][ dy ][ dz ] = tmpIdx  
 if( cellDoSmoothing[ tmpIdx ] && cellCnt[ tmpIdx ] )  
 otherClusterPtCnt = 1  
 }

### Trilinear Filtering

Inputs to this process are:

* gridSize,
* numBoundaryCells, the number of cells containing a boundary point or neighboring such a cell,
* pos3D[ n ], n =0..2, position of a reconstructed point,
* centroid[ x ][ y ][ z ][ k ], for x, y, and z in the range of 0 to 1, inclusive, and k =0..2, an array storing the centroids on the 2 × 2 ×2 grid of cells,
* cellAnchor[ k ], k = 0..1, an array storing the index of anchor of the 2 × 2 ×2 grid of cells in x, y, and z directions, respectively,
* idxArr[ dx ][ dy ][ dz ], dx, dy, and dz in the range of 0 to 1, inclusive, an array storing the 1-D indices for the 2 × 2 ×2 grid of cells, and
* cellCnt[ n ], n = 0..numBoundaryCells – 1, an array containing number of reconstructed points in a cell.

Output of this process is:

* pos3dFilt[ k ], k = 0..1, a filtered version of the reconstructed point pos3D[ k ], k =0..2.

[Ed. (RJ): This needs to be carefully extended so that it may be used for attribute smoothing as well.]

This process carries out trilinear filtering for geometry smoothing. The following steps apply:

* The weights for trilinear filtering are calculated as follows:

for( k = 0; k < 3 ; k++ ) {  
 w[ k ][ 1 ] = ( pos3D[ k ] – ( cellAnchor[ k ] \* gridSize) – ( gridSize / 2) ) \* 2 + 1  
 w[ k ][ 0 ] = (gridSize \* 2) – w[ k ][ 1 ]  
 }

* The trilinear filtering is performed as follows:

cnt= 0  
for( k = 0; k < 3; ++)  
 filtVal[ k ] = 0  
for( dx = 0; dx < 2; dx++ ) {  
 wx = w[ 0 ][ dx ]  
 for( dy = 0; dy < 2; dy++ ) {  
 wy = w[ 1 ][ dy ]  
 for( dz = 0; dz < 2; dz++ ) {  
 wz = w[ 2 ][ dz ]  
 tmpIdx = idxArr[ dx ][ dy ][ dz ]  
 for( k = 0; k < 3; ++) {  
 filtVal[ k ] += wx \* wy \* wz \* centroid [dx ][ dy ][ dz ][ k ]  
 cnt += wx \* wy \* wz \* cellCnt[ tmpIdx ]  
 }  
 }  
 }  
}  
cnt = cnt / ( 8 \* gridSize \* gridSize \* gridSize )  
for( k = 0; k < 3; ++) {  
 filtOut[ k ] = filtOut[ k ] ÷ ( 8 \* gridSize \* gridSize \* gridSize )  
 pos3dFilt[ k ] = filtOut[ k ] \* cnt  
}

### Patch Border Filtering process

This process is invoked when a bitstream contains a geometry smoothing SEI message with a value of gs\_smoothing\_method\_type[ k ] equal to 2, and the point cloud reconstruction system has selected to use this k-th method for performing smoothing during point cloud reconstruction. Patch border filtering is a process acting upon the occupancy map.

#### Patch Border Filtering

Inputs to this process are:

* OFrame[ y ][ x ], the occupancy map corresponding to the current point cloud frame at nominal resolution, where y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1.
* GFrame[ 0 ][ y ][ x ], the first map of the decoded geometry frames at nominal resolution, where y = 0..asps\_frame\_height – 1, x = 0.. asps\_frame\_width – 1.
* the index of the patch, pIdx.
* PbfThreshold[ k ], PbfPasses[ k ], PbfSizeX[ k ] and PbfSizeY[ k ] as defined in the semantics of the geometry smoothing SEI message.

Output to this process is:

* updated OFrame[ y ][ x ], the updated occupancy map corresponding to the current point cloud frame at nominal resolution, where y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1.

First, boundary points of patches and neighbouring patches are determined as follows:

* Clause 9.6.5 is invoked with OFrame[ y ][ x ] and pIdx as inputs. The output is a container recPCboundaryPointType holding the list of patch boundary type for the reconstructed point cloud.

[Ed. (PA): This needs to be changed according to 9.6.5. Need to understand what PointToPixel correspond to and how applying it to PBF]

* For pIdx = 0 .. AtlasTotalNumberOfPatches[ OrdIdx ] − 1, the following applies:
  + For u = 0..Patch2dSizeX[ pIdx ] − 1, v = 0..Patch2dSizeY[ pIdx ] – 1, the following steps apply:
    - The conversion of patch coordinates to atlas coordinates as specified in clause 9.4.5 is invoked with patch coordinates (u, v), and patch index pIdx, as inputs and atlas coordinates (x, y) as output.
    - The 3D coordinates corresponding to the point from recPCboundaryPointType that is marked as located near a patch boundary are derived by invoking clause 9.4.6 with ( u, v ), pIdx, and GFrame[ 0 ][ y ][ x ] as inputs. The output is a vector, pos3D of size 3, containing the (x, y, z) coordinates of the point.
    - The point is added to the 3D boundary points list borderPoints[ pIdx ] of size borderPointsSize[ pIdx ] and the border bounding box borderBoundingBox[ pIdx ] is updated by invoking clause 9.6.7.3 with the (x, y, z) coordinates of the added point as input and updated borderBoundingBox[ pIdx ] as output.
* For pIdx = 0 .. AtlasTotalNumberOfPatches[ OrdIdx ] − 1, the following applies:
  + Determine the patches different from patch pIdx which bounding box intersects patch pIdx bounding box and add these patches index to neighbouringPatches[ pIdx ] of size neighPatchesSize[ pIdx ], the list of neighbouring patches for a patch of index pIdx.

Then, the depth of the 3D border points of the boundary patches of patch pIdx is determined as follows:

* For pIdx = 0.. AtlasTotalNumberOfPatches[ OrdIdx ] − 1, q = 0..neighPatchesSize[ pIdx ] − 1,  
  b = 0..borderPointsSize[ q ]
  + if b belongs to borderBoundingBox[ pIdx ], the following ordered steps apply:
    - The conversion of the coordinates of 3D point position to patch coordinates and depth as specified in clause 9.6.7.2 is invoked with 3D point b of coordinates stored in vector pos of size 3 projected on patch of index pIdx as inputs and patch coordinates (u, v) and depth borderPointDepth as output.
    - The conversion of patch coordinates to atlas coordinates as specified in clause 9.4.5 is invoked with patch coordinates (u, v), and patch index pIdx, as inputs and atlas coordinates (x, y) as output.
    - if abs( borderPointDepth - GFrame[ 0 ][ y ][ x ] ) < PbfThreshold[ k ]
      * neighbourDepth[ pIdx ][ y ][ x ] = borderPointDepth

Then, the occupancy map is eroded as specified by the following steps:

* An image oPbfFrame is copied from OFrame and initialized to 0
* For pIdx = 0.. AtlasTotalNumberOfPatches[ OrdIdx ] − 1, iter = 0..PbfPasses[ k ] − 1, u = 0..Patch2dSizeX[ pIdx ] − 1, v = 0..Patch2dSizeY[ pIdx ] – 1, the following applies:
  + The conversion of patch coordinates to atlas coordinates as specified in clause 9.4.5 is invoked with patch coordinates (u, v), and patch index pIdx, as inputs and atlas coordinates (x, y) as output.
  + If OFrame[ y ][ x ] is not equal to 0, the following steps apply:
    - The variable countNeighbour is derived as follows:

countNeighbor = OFrame[ y ][ x − 1 ] + OFrame[ y ][ x + 1 ]  
 + OFrame[ y − 1 ][ x ] + OFrame[ y + 1 ][ x ]

* + - The smoothing of the occupancy map applies as follows:
      * if countNeighbour is equal to 0, oPbfFrame[ y ][ x ] is set equal to 0
      * else if countNeighbour is less than 4, the smoothing of oPbfFrame[ y ][ x ] as specified in clause 9.6.7.4 is invoked with OFrame[ y ][ x ], GFrame[ 0 ][ y ][ x ], oPbfFrame[ y ][ x ], neighbourDepth[ pIdx ][ y ][ x ], PbfSizeX[ k ] and PbfSizeY[ k ] as inputs and updated oPbfFrame[ y ][ x ] as output.
      * else oPbfFrame[ y ][ x ] is set equal to 1.

As a final stage, the computed occupancy map oPbFrame overwrites OFrame.

#### Conversion from 3D point position to patch coordinates and depth

The inputs to this process are:

* a vector, pos3D, of dimension 3,
* patch index pIdx.

The outputs of this process are:

* patch coordinates (u, v) ,
* depth value depthValue.

The following steps apply:

u = (pos3D[ PatchAxisX[ pIdx ]] – Patch3dPosX[ pIdx ]) / PatchLoDScaleX[ pIdx ]

v = (pos3D[ PatchAxisY[ pIdx ]] – Patch3dPosY[ pIdx ]) / PatchLoDScaleY[ pIdx ]

projectionScale = (1 – 2\* PatchProjectionMode[ pIdx ])

depthValue = (pos3D[ PatchAxisZ[ pIdx ] ] - Patch3dPosMinZ[ pIdx ]) \* projectionScale

#### Resize 3D bounding box

The input of this process is:

* a vector, pos3D, of dimension 3.

The output of this process is:

* updated bounding box, boundingBox

The bounding box is a cubic space in 3D defined by two 3D points: boundingBoxMin[3] and boundingBoxMax[3] which values are respectively initialized to: {+inf; +inf; + inf} and {0; 0; 0}.

One point is added to the bounding box by updating the positions of the two bounding box points defining the bounding box with the following steps:

    if ( boundingBoxMin[0] > pos3D[0] ) boundingBoxMin[0] = pos3D[0]

    if ( boundingBoxMin[1] > pos3D[1] ) boundingBoxMin[1] = pos3D[1]

    if ( boundingBoxMin[2] > pos3D[2] ) boundingBoxMin[2] = pos3D[2]

    if ( boundingBoxMax[0] < pos3D[0] ) boundingBoxMax[0] = pos3D[0]

    if ( boundingBoxMax[1] < pos3D[1] ) boundingBoxMax[1] = pos3D[1]

    if ( boundingBoxMax[2] < pos3D[2] ) boundingBoxMax[2] = pos3D[2]

#### Occupancy map smoothing

The inputs of this process are:

* OFrame[ y ][ x ], the occupancy map corresponding to the current point cloud frame at nominal resolution, where y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1.
* GFrame[ 0 ][ y ][ x ], the first map of the decoded geometry frames at nominal resolution, where y = 0..asps\_frame\_height – 1, x = 0.. asps\_frame\_width – 1.
* oPbfFrame[ y ][ x ], a temporary occupancy map, where y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1.
* neighbourDepth[ pIdx ][ y ][ x ]
* PbfSizeX[ k ] and PbfSizeY[ k ]

The output of this process is:

* updated oPbfFrame[ y ][ x ], where y = 0..asps\_frame\_height – 1, x = 0..asps\_frame\_width – 1.

The following steps apply:

* The variable boundaryOrientation is derived as follows:

boundaryOrientation = orientationArray[ (OFrame[ y − 1 ][ x − 1 ]<<7) |

(OFrame[ y − 1 ][ x ]<<6) | OFrame[ y − 1 ][ x + 1 ]<<5) | (OFrame[ y ][ x − 1]<<4) |

(OFrame[ y ][ x+1 ]<<3) |(OFrame[ y + 1 ] [ x − 1 ]<<2) | (OFrame[ y + 1 ][ x ]<<1) |

(OFrame[ y + 1][ x + 1 ]) ]

where the array orientationArray is defined as follows:

orientationArray = { 0,0,6,0,0,0,0,6,4,0,0,5,0,0,0,5,0,0,0,0,0,0,7,7,0,0,0,0,0,0,0,6,0,0,0,0,0,0,0,0,0,4,0,5,0,0,0,5,

0,0,0,0,0,0,7,0,0,0,0,0,0,0,0,5,2,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,

0,0,0,0,0,0,0,0,3,3,0,4,3,0,0,5,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,5,0,0,0,0,

0,0,0,0,0,0,7,7,0,0,0,0,0,0,0,7,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,6,

0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,1,0,0,7,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,3,3,0,3,0,0,0,4,

1,0,0,0,1,0,1,0,2,3,0,0,1,2,0,0}

* The vectors vX and vY are derived as follows:

vX = vectorOrientation[ orientationArray ]

vY = vectorOrientation[ (orientationArray + 2) % 8 ]

where vectorOrientation is defined as follows:

vectorOrientation[ 8 ][ 2 ] ={ {1,0}, {1,1}, {0,1}, {-1,1}, {-1,0}, {-1,-1}, {0,-1}, {1,-1} }

* The variables sumP, sumE and windowCount are initialized to 0
* For xW = − PbfSizeX[ k ]..PbfSizeX[ k ], yW = − PbfSizeY[ k ]..PbfSizeY[ k ]

if( neighbourDepth[ pIdx ][ y + vX[ 1 ]\*xW + vY[ 1 ]\*yW ][ x + vX[ 0 ]\*xW + vY[ 0 ]\*yW ] )

* + - sumP += ( xW\*vX[ 0 ]+yW\*vY[ 0 ] )2+ ( xW\*vX[ 1 ] + W\*vY[ 1 ])2 +( neighbourDepth[ pIdx ][ y + vX[ 1 ] \* xW + vY[ 1 ] \* yW][x + vX[ 0 ] \* xW  
      + vY[ 0 ] \* yW ] − GFrame[ 0 ][ y ][ x ] )2
    - sumE += ( (xW + 1)\*vX[ 0 ]+yW\*vY[ 0 ] )2+ (xW + 1)\*vX[ 1 ] + yW\*vY[ 1 ])2  
      +( neighbourDepth[ pIdx ][ y + vX[ 1 ] \* xW + vY[ 1 ] \*yW][x + vX[ 0 ] \* xW   
      + vY[ 0 ] \* yW ] − GFrame[ 0 ][ y − vX[ 1] ][ x − vX[ 0 ] ] )2
    - The variable windowCount is incremented by 1
* oPbfFrame[ y ][ x ] = ( windowCount == 0 ) || ( sumE >= sumP )

# Parsing process

## General

Inputs to this process are bits from the bitstream

Outputs of this process are syntax element values.

This process is invoked when the descriptor of a syntax element in the syntax tables is equal to ue(v) or se(v) (see clause 10.2).

## Parsing process for 0-th order Exp-Golomb codes

### General

This process is invoked when the descriptor of a syntax element in the syntax tables is equal to ue(v) or se(v).

Inputs to this process are bits from the bitstream.

Outputs of this process are syntax element values.

Syntax elements coded as ue(v) or se(v) are Exp-Golomb-coded. The parsing process for these syntax elements begins with reading the bits starting at the current location in the bitstream up to and including the first non-zero bit and counting the number of leading bits that are equal to 0. This process is specified as follows:

leadingZeroBits = −1  
for( b = 0; !b; leadingZeroBits++ ) (10‑1)  
 b = read\_bits( 1 )

The variable codeNum is then assigned as follows:

codeNum = 2leadingZeroBits − 1 + read\_bits( leadingZeroBits ) (10‑2)

where the value returned from read\_bits( leadingZeroBits ) is interpreted as a binary representation of an unsigned integer with the most significant bit written first.

Table 10‑1 illustrates the structure of the Exp-Golomb code by separating the bit string into "prefix" and "suffix" bits. The "prefix" bits are those bits that are parsed as specified above for the computation of leadingZeroBits and are shown as either 0 or 1 in the bit string column of Table 10‑1. The "suffix" bits are those bits that are parsed in the computation of codeNum and are shown as xi in Table 10‑1, with i in the range of 0 to leadingZeroBits − 1, inclusive. Each xi is equal to either 0 or 1.

Table 10‑1 Bit strings with "prefix" and "suffix" bits and assignment to codeNum ranges (informative)

|  |  |
| --- | --- |
| **Bit string form** | **Range of codeNum** |
| 1 | 0 |
| 0 1 x0 | 1..2 |
| 0 0 1 x1 x0 | 3..6 |
| 0 0 0 1 x2 x1 x0 | 7..14 |
| 0 0 0 0 1 x3 x2 x1 x0 | 15..30 |
| 0 0 0 0 0 1 x4 x3 x2 x1 x0 | 31..62 |
| ... | ... |

Table 10‑2 illustrates explicitly the assignment of bit strings to codeNum values.

Table 10‑2 – Exp-Golomb bit strings and codeNum in explicit form and used as ue(v) (informative)

|  |  |
| --- | --- |
| **Bit string** | **codeNum** |
| 1 | 0 |
| 0 1 0 | 1 |
| 0 1 1 | 2 |
| 0 0 1 0 0 | 3 |
| 0 0 1 0 1 | 4 |
| 0 0 1 1 0 | 5 |
| 0 0 1 1 1 | 6 |
| 0 0 0 1 0 0 0 | 7 |
| 0 0 0 1 0 0 1 | 8 |
| 0 0 0 1 0 1 0 | 9 |
| ... | ... |

Depending on the descriptor, the value of a syntax element is derived as follows:

– If the syntax element is coded as ue(v), the value of the syntax element is equal to codeNum.

– Otherwise (the syntax element is coded as se(v)), the value of the syntax element is derived by invoking the mapping process for signed Exp-Golomb codes as specified in clause 10.2.2 with codeNum as input.

### Mapping process for signed Exp-Golomb codes

Input to this process is codeNum as specified in clause 10.2.1.

Output of this process is a value of a syntax element coded as se(v).

The syntax element is assigned to the codeNum by ordering the syntax element by its absolute value in increasing order and representing the positive value for a given absolute value with the lower codeNum. Table 10‑3 provides the assignment rule.

Table 10‑3 – Assignment of syntax element to codeNum   
for signed Exp-Golomb coded syntax elements se(v)

|  |  |
| --- | --- |
| **codeNum** | **syntax element value** |
| 0 | 0 |
| 1 | 1 |
| 2 | −1 |
| 3 | 2 |
| 4 | −2 |
| 5 | 3 |
| 6 | −3 |
| k | (−1)k + 1 Ceil( k ÷ 2 ) |

[Ed. (RJ) (AMT): Should we use (k+1)/2 or (k+1)>>1 instead of Ceil( k ÷ 2)?]

1. Profiles, tiers, and levels
   1. Overview of profiles, tiers, and levels

Profiles, tiers, and levels specify restrictions on the bitstreams and hence limits on the capabilities needed to decode the bitstreams. Profiles, tiers, and levels may also be used to indicate interoperability points between individual decoder implementations.

Each profile specifies a subset of algorithmic features and limits that shall be supported by all decoders conforming to that profile.

Each level of a tier specifies a set of limits on the values that may be taken by the syntax elements of this document. The same set of tier and level definitions is used with all profiles, but individual implementations may support a different tier and within a tier a different level for each supported profile. For any given profile, a level of a tier generally corresponds to a particular decoder processing load and memory capability.

Capabilities of V-PCC decoders conforming to this document are specified in terms of the ability to decode V-PCC streams conforming to the constraints of profiles, tiers and levels specified in this annex and other annexes. When expressing the capabilities of a decoder for a specified profile, the tier and level supported for that profile should also be expressed.

NOTE – The term “decode” in this context may or may not include the reconstruction in 3D space, i.e. “decode” may refer to only decoding the 2D video sub-bitstreams (i.e. geometry, occupancy, attribute(s)) and the associated atlas sub-bitstream, or it may include further reconstruction into 3D space. The respective steps included in the “decoding” are specified for each profile.

* 1. V-PCC profile, tier and level structure

V-PCC profiles follow a structured and flexible definition to allow for clearly identifying two distinct conformance points. These conformance points are illustrated in the decoding block diagram shown in Figure A-1.

The first conformance point, point A, covers the decoded attributes, geometry, and occupancy video sub-bitstreams, plus the decoded atlas sub-bitstream. It also covers the derived block to patch map information. It does not, however, cover the point cloud reconstruction process.

The second conformance point, point B, covers the point cloud reconstruction process.

A V-PCC profile consists of a combination of up to three profile components, identified from the syntax elements ptl\_profile\_codec\_group\_idc, ptl\_profile\_pcc\_toolset\_idc, and ptl\_profile\_reconstruction\_idc, as shown in Figure A-2.

The first two profile components together describe conformance point A. More specifically, they describe the supported:

* video encoding specifications and their profiles (e.g., AVC Progressive High, HEVC Main/Main 10 etc), referred to as the **CodecGroup** profile component, and
* V-PCC specific tools (e.g., use of EOM and PLR etc), referred to as the **Toolset** profile component. The **Toolset** profile component describes the bitstream syntax structure.

[Ed. Note (all): Conformance to point A should also include the relationships between the sub-bitstreams that are decoded. For example, conformance point A should, at a minimum, include atlas, occupancy, and geometry map 0 sub-bitstreams. Additional geometry maps may be included. Attributes, could also be included (with no specific order), but the attributes should also be associated with a corresponding geometry map. An attribute map that does not have an associated geometry map should be discarded.

Furthermore, conformance can be achieved using sub-bitstreams and not necessarily through the decoding of a V-PCC elementary stream.].

[Ed. Note (SS): Regarding note above, I added some text under the toolset clause.].

The third profile component describes conformance point B, specifying the reconstruction tools supported or recommended to achieve conformance in 3D reconstruction. This is referred to as the **Reconstruction** profile component.

[Ed. Note (SS): What has to be done here?]

Together these profile components form the V-PCC profiles, following the naming convention **V-PCC CodecGroup Toolset Reconstruction**. For example, taking the first blocks of Figure A-2, the resulting V-PCC profile is named **V-PCC AVC Basic Rec0*.***



Figure A-1: V-PCC decoding block diagram with decoding conformance points A and B.



**Figure A-2: The V-PCC profile, tier, and level structure.**

Conformance can be specified for a single V-PCC bitstream, as specified in Annex B, or can be a collection of V-PCC sub-bitstream components that together specify a minimal collection of sub-bitstreams needed to reconstruct a point cloud. The minimal collection shall include the atlas, occupancy, and geometry map with index 0 sub-bitstreams. Additional geometry map sub-bitstreams should also be included. Attributes could also be included (with no specific order), but the included attributes shall be associated with a corresponding geometry map. An attribute map of index i that does not have an associated geometry map with the same index i should be discarded.

Any V-PCC bitstream, collection of V-PCC sub-bitstream components, or V-PCC decoder can claim conformance at two points. However, indicating conformance point A is mandatory, while for conformance point B it is optional, based on the selected profile. For example:

* **V-PCC AVC Basic,** is a valid profile to indicatedecodingwithout any 3D reconstruction.
* **V-PCC AVC Basic Rec0,** is a valid profile to indicatedecodingwith elementary 3D reconstruction.
* **V-PCC AVC Rec0,** is NOT a valid profile as the Toolset information is missing.
* **V-PCC Basic Rec0,** is NOT a valid profile as the CodecGroup information is missing.

To indicate profile conformance of a bitstream at conformance point A, it is needed to signal the CodecGroup and Toolset profile components. To indicate profile conformance of a bitstream at conformance point B, it is needed to use all three profile components.

[Ed Note (AMT): My interpretation of the Reconstruction profile is that this is an indication of a preferred reconstruction process that a decoder, if it supports it, should (not shall), consider to use. A decoder may in fact support multiple reconstruction profiles and it may select, based on other conditions, e.g. power restrictions, to follow a different reconstruction strategy. Conformance to a reconstruction profile is only a decoder requirement, not a bitstream requirement. If a decoder supports a particular reconstruction profile (and all decoders should support Rec.0) and if it selects to use when decoding a bitstream that particular reconstruction profile, then it shall follow the process specified by that profile. However, it may still select to use another reconstruction process, and for that instance it may conform to another process. Others may disagree with this interpretation. ]

Decoders conforming to a V-PCC profile at conformance point A (identified by syntax elements ptl\_profile\_codec\_group\_idc and ptl\_profile\_pcc\_toolset\_idc) at a specific level (identified by a specific value of syntax element ptl\_level\_idc) of a specific tier (identified by a specific value of syntax element ptl\_tier\_flag) shall be capable of decoding all V-PCC bitstreams or collection of V-PCC sub-bitstream components, according to clause 8 (Decoding process), for which all of the following conditions apply:

* The V-PCC bitstream or the collection of V-PCC sub-bitstream components are indicated to conform to the supported CodecGroup and Toolset profile components, as indicated in clauses A.3 (CodecGroup) and A.4 (Toolset).
* The V-PCC bitstream or the collection of V-PCC sub-bitstream components are indicated to conform to a level that is lower than or equal to the specified level, as indicated in clause A.6.
* The V-PCC bitstream or the collection of V-PCC sub-bitstream components are indicated to conform to a tier that is lower than or equal to the specified tier, as indicated in clause A.6.

[Ed Note (AMT): The below text is debatable since it places restrictions on a decoder given reconstruction. It does not provide an option to a decoder to maybe select a different reconstruction, if it supports multiple. In my opinion reconstruction can be selected, not enforced. Others may disagree]

Decoders conforming to a V-PCC profile at conformance point B (identified by syntax elements ptl\_profile\_codec\_group\_idc, ptl\_profile\_pcc\_toolset\_idc, and ptl\_profile\_reconstruction\_idc) at a specific level (identified by a specific value of ptl\_level\_idc) of a specific tier (identified by a specific value of ptl\_tier\_flag) shall be capable of decoding all PCC bitstreams, according clause 8 (Decoding process) and clause 9 (Reconstruction process) for which all of the following conditions apply:

* The V-PCC bitstream or the collection of V-PCC sub-bitstream components are is indicated to conform to the supported CodecGroup, Toolset, and Reconstruction profile type.
* The V-PCC bitstream or the collection of V-PCC sub-bitstream components are indicated to conform to a level that is lower than or equal to the specified level.
* The V-PCC bitstream or the collection of V-PCC sub-bitstream components are indicated to conform to a tier that is lower than or equal to the specified tier.
  1. V-PCC CodecGroup profile components

Table A-1 provides a list of the available CodecGroup profile components for V-PCC and respectively the allowed values for syntax element ptl\_profile\_codec\_group\_idc.

**Table A-1 – Available CodecGroup profiles components**

|  |  |
| --- | --- |
| **CodecGroup** | **ptl\_profile\_codec\_group\_idc** |
| AVC Progressive High | 0 |
| HEVC Main10 | 1 |
| HEVC444 | 2 |
| MP4RA | 3 |

Table A-2 provides a list of the supported codec functionalities for each CodecGroup profile component.

**Table A-2 – CodecGroup profile component supported functionality**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | AVC Progressive High | | | HEVC Main10 | | | HEVC444 | | | MP4RA | | |
|  | Capability | Occupancy | Geometry | Attributes | Occupancy | Geometry | Attributes | Occupancy | Geometry | Attributes | Occupancy | Geometry | Attributes |
| Chroma format | Mono |  |  |  |  |  |  | ✓ | ✓ | ✓ | ‒ | ‒ | ‒ |
| 4:2:0 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ‒ | ‒ | ‒ |
| 4:4:4 |  |  |  |  |  |  |  |  | ✓ | ‒ | ‒ | ‒ |
| Bitdepth | 8 bit | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ‒ | ‒ | ‒ |
| 10 bit |  |  |  | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ‒ | ‒ | ‒ |
| Video coding specification | name | ISO/IEC 14496-10(AVC) | | | ISO/IEC 23008-2 (HEVC) | | | | | | As defined by a component codec mapping SEI (E.2.13) | | |
| profile | Progressive High | | | Main 10 | | | Main 4:4:4 10 | | | ‒ | | |

A tick mark in the above Table indicates that a particular feature is supported by the defined profile. A dash symbol, i.e. "‒", indicates that support of a particular feature is specified by means outside this document.

All video sub-bitstreams of a V-PCC bitstream or a collection of V-PCC sub-bitstreams conform to the stream format indicated in the component codec mapping SEI (E.2.13).

The use of the MP4RA CodecGroup profile component is not meant as an interoperability point. Conformance to this profile can be specified by means outside this document.

* 1. V-PCC toolset profile components

Table A-3 provides a list of the available toolset profile components for V-PCC and respectively the allowed values for syntax element ptl\_profile\_toolset\_idc.

**Table A-3 – Available toolset profile components**

|  |  |
| --- | --- |
| **Toolset** | **ptl\_profile\_pcc\_toolset\_idc** |
| Basic | 0 |
| Extended | 1 |

* + 1. Basic and Extended toolset profile component

V-PCC toolset profile components indicating the Basic (ptl\_profile\_pcc\_toolset\_idc = 0) or the Extended (ptl\_profile\_pcc\_toolset\_idc = 1) toolset profile component shall conform to the syntax element restrictions specified in Table A-4. If a syntax element is not mentioned in Table A-4,, it is not restricted through a toolset profile component.

**Table A-4– Required syntax element values for basic and extended toolset profile components**

|  |  |  |
| --- | --- | --- |
|  | **Max allowed values** | |
|  | **ptl\_profile\_pcc\_toolset\_idc** | |
| **Syntax element** | **0 (basic)** | **1 (extended)** |
| asps\_eom\_patch\_enabled\_flag | 0 | ‒ |
| asps\_map\_count\_minus1 | min( 1, LevelMapCount – 1 ) | LevelMapCount – 1 |
| vps\_multiple\_map\_streams\_present\_flag  (when vps\_map\_count\_minus1 > 0) | when present, 1 | ‒ |
| vps\_atlas\_count\_minus1 | 0 | 0 |
| asps\_point\_local\_reconstruction\_enabled\_flag | 0 | ‒ |
| ai\_attribute\_dimension\_minus1 | 2 | ‒ |
| ai\_attribute\_dimension\_partitions\_minus1 | 0 | ‒ |
| ai\_attribute\_partition\_channels\_minus1 | ‒ | 2 |
| asps\_use\_eight\_orientations\_flag | 0 | ‒ |
| asps\_extended\_projection\_enabled\_flag | 0 | ‒ |

Furthermore, the following restrictions shall apply to a bitstream or a collection of V-PCC sub-bitstream components conforming to the Basic toolset profile components:

* At a minimum, there should be an atlas, a geometry video, and an occupancy video sub-bitstream present. [Ed. (JB): Should refer to an output unit.]
* For each ordered output picture orderIdx there shall be at least present
  + decOccFrame[ orderIdx ][ compIdx ][ y ][ x ], and its associated bitdepth, decOccBitdepth[ orderIdx ], luma width, decOccWidth[ orderIdx ], and luma height, decOccHeight[ orderIdx ], where orderIdx is the display output order index of the decoded occupancy frames, compIdx corresponds to the colour component index and is equal to 0, y is the row index in the decoded frame and is in the range of 0 to decOccHeight[ orderIdx ] − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to decOccWidth[ orderIdx ] − 1, inclusive, as specified in clause 8.2.
  + decGeoFrame[ mapIdx ][ orderIdx ][ compIdx ][ y ][ x ], and its associated bitdepth, decGeoBitdepth[ mapIdx ][ orderIdx ], luma width, decGeoWidth[ mapIdx ][ orderIdx ], and luma height, decGeoHeight[ mapIdx ][ orderIdx ], where mapIdx corresponds to the map index of each frame and is in the range of 0 to vps\_multiple\_map\_streams\_present\_flag[ atlasIdx ] ? vps\_map\_count\_minus1[ atlasIdx ] : 0, inclusive, orderIdx is the display output order index of the decoded geometry frames, compIdx corresponds to the colour component index and is equal to 0, y is the row index in the decoded frame and is in the range of 0 to decGeoHeight[ mapIdx ][ orderIdx ] − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to decGeoWidth[ mapIdx ][ orderIdx ] − 1, inclusive, as specified in clause 8.3.
  + parsed syntax structures encapsulated within atlas data unit [orderIdx], where orderIdx is the displayoutput order index of the decoded atlas frames, as specified in clause 8.5

[Ed Note (SS): Need to define output unit]

* The display output order index orderIdx of the decoded output units shall be the same for decOccFrame, decGeoFrame, atlas data unit, and any present decAttrFrame[ attrIdx ].

[Ed. (JB): Should refer to an output unit.]

* All video sub-bitstreams shall have an intra random access point (IRAP) at IRAP points of an atlas sub-bitstream.

[Ed. (JB): This should be expressed elsewhere. Should define an IRAP access unit, which requires IRAP pictures in a video sub-bitstreams.]

* If syntax element asps\_pixel\_deinterleaving\_enabled\_flag or asps\_point\_local\_reconstruction\_enabled\_flag are equal to 1, asps\_map\_count\_minus1 shall be equal to 0.
  1. V-PCC reconstruction profile components

Table A-6 provides a list of the available reconstruction profile components for V-PCC and respectively the allowed values for syntax element ptl\_profile\_reconstruction\_idc.

**Table A-5 – Available reconstruction profile components**

|  |  |
| --- | --- |
| **reconstruction profile component** | **ptl\_profile\_reconstruction\_idc** |
| Rec0 | 0 |
| Rec1 | 1 |
| Rec2 | 2 |
| Rec Unconstrained | 3 |

[Ed (PA): Rec Unconstrained would be better with allocated value 0 in case other Rec appear later]

Table A-6 provides a list of the supported reconstruction functionalities for each reconstruction profile component.

**Table A-6 – Supported reconstruction operations for each reconstruction profile component**

|  |  |  |  |
| --- | --- | --- | --- |
| **Reconstruction Operation** | **Rec0** | **Rec1 and Rec2** | **Rec Unconstrained** |
| chroma format conversion | see clause A.5.1.1 | see clause A.5.1.1 |  |
| resolution upsampling | see clause A.5.1.2 | see clause A.5.1.2 |  |
| Frame rate conversion | see clause A.5.1.3 | see clause A.5.1.3 |  |
| Pixel deinterleaving | ignored | reconstructed,  according to clause 9.4.1 | unconstrained |
| PLR | ignored | reconstructed,  according to clause 9.4.2 | unconstrained |
| EOM | ignored | reconstructed,  according to clause 9.4.4 | unconstrained |
| Duplicate point removal | ignored | reconstructed,  according to clause 9.4.8 | unconstrained |
| RAW patches | ignored | reconstructed,  according to clause 9.5 | unconstrained |
| Smoothing | ignored | reconstructed,  according to clause 9.6 | unconstrained |

[Ed (PA): Rec2 only differs from Rec1 with PBF on and Grid-based smoothing (GS) off and reciprocally. There are current discussions about having PBF in OCM SEI, like we have GS in an GEO SEI and CS (colour smoothing) in ATTR SEI. As a consequence, we might have only a Rec0 and Rec1 profiles?... Besides there is discussion as well on the naming of such profile Rec0, Rec1 like we get Basic, Extended for CPA etc...]

* + 1. V-PCC reconstruction decoder conformance

Decoders conforming to a V-PCC profile with a reconstruction profile component must perform reconstruction operations. Conformance is assessed at conformance point A (decoded attribute, geometry, and occupancy bitstreams together with decoded patch metadata and decoded block to patch map) and conformance point B (reconstructed point cloud).

In order to check for decoder conformance at point B the following hypothetical reference operations are defined:

* Hypothetical reference chroma format conversion (A.5.1.1)
* Hypothetical reference resolution upsampling (A.5.1.2)
* Hypothetical reference frame rate conversion (A.5.1.3)
* Hypothetical reference attribute transfer process (A.5.1.4)

A decoder may wish to use these processes exactly as defined in their corresponding clauses. However, a decoder can still be considered as conforming to this document if the resulting point cloud is perceived as similar in reconstruction as the point cloud constructed using these hypothetical reference processes. Similarity of these processes with the hypothetical reference processes is outside of the scope of this document.

For decoder conformance of the attribute values associated with each point in the point cloud, no further conversion to any other domain other than the domain used by the video representation associated with the respective attribute, shall be used. For example, if an attribute of type ATTR\_TEXTURE with three components was coded using the HEVC specification and was associated with parameters colour\_primaries equal to 1, transfer\_characteristics equal to 1, and matrix\_coeffs equal to 1, then the attribute is assumed to be in the ITU-R BT.709 non-constant luminance YCbCr representation.

* + - 1. Hypothetical reference chroma format conversion

When an attribute is of type ATTR\_TEXTURE and has three components with a chroma format other than 4:4:4, it is converted to a 4:4:4 chroma format according to its specified chroma sample location. For conformance purposes, the 4 tap filter specified in Table A-7 shall be used:

**Table A-7 – 4-phase chroma resampling filter**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Phase p** | **Interpolation coefficients** | | | |
| **fc[ p, 0 ]** | **fc[ p, 1 ]** | **fc[ p, 2 ]** | **fc[ p, 3 ]** |
| 0 | 0 | 32 | 0 | 0 |
| 1 | −1 | 8 | 27 | −2 |
| 2 | −2 | 18 | 18 | −2 |
| 3 | −2 | 27 | 8 | −1 |

* + - 1. Hypothetical reference resolution upscaling

For the occupancy map, pixel replication is used to upsample the occupancy map to a width of asps\_nominal\_width and a height of asps\_nominal\_height.

For all attributes except those of type ATTR\_TEXTURE, pixel replication is used to upsample the attribute to a width of asps\_nominal\_width and a height of asps\_nominal\_height.

For attributes of type ATTR\_TEXTURE, the upsampling filters specified in clause H.8.1.4.2.2, Table H.1 of the HEVC specification are used to upsample the attribute to a width of asps\_nominal\_width and a height of asps\_nominal\_height.

For geometry, bilinear interpolation is used to upsample the geometry to a width of asps\_nominal\_width and a height of asps\_nominal\_height.

* + - 1. Hypothetical reference frame rate conversion

Frame repetition using the nearest available frame from the past is used to perform frame rate conversion.

* + - 1. Hypothetical reference attribute transfer process

The inputs to this process are:

* pointCnt, the number of points in the current reconstructed point cloud frame,
* isBoundaryPoint[ n ], n = 0..pointCnt – 1, indicating whether the n-th reconstructed point lies on or near a patch boundary, and
* recPcGeoSmoothed[ i ][ k ], i = 0..pointCnt – 1, k = 0..2, an array containing the reconstructed point positions, and
* recPcAttr[ ]

Output of this process are:

* recPcAttrMod[ ]

This process assigns new attribute values for reconstructed points that have been modified by the geometry smoothing process. The following applies:

* For n = 0..pointCnt – 1, the following applies:
* If isBoundaryPoint[ n ] is equal to 2, the following attribute transfer process is applied.
  1. Tiers and Levels

For purposes of comparison of tier capabilities, the tier with ptl\_tier\_flag equal to 0 is considered to be a lower tier than the tier with ptl\_tier\_flag equal to 1. Currently only a single tier, the Main tier, is specified for V-PCC. This is indicated by setting the syntax element ptl\_tier\_flag equal to 0. It is a requirement for bitstream conformance to this current version of the Specification that ptl\_tier\_flag shall always be equal to 0.

For purposes of comparison of level capabilities, a particular level of a specific tier is considered to be a lower level than some other level of the same tier when the value of the ptl\_level\_idc of the particular level is less than that of the other level.

Table A-8 specifies the general point cloud and VPS related limits for each level of each tier.

Table A-9 specifies the general atlas ASPS and tile related limits for each level of each tier.

Table A-10 specifies the general video bitstream related limits for each level of each tier.

A tier and level to which a bitstream conforms are indicated by the syntax elements ptl\_tier\_flag and ptl\_level\_idc. ptl\_level\_idc shall be set equal to a value of 30 times the level number specified in Table A-8, Table A-9, and Table A-10.

Table A-8 — General Point Cloud or VPS related level limits

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Level** | **Max # of points per second**  **MaxNumPointsPerSec** | **Max # of points per frame**  **MaxNumPointsPerFrame** | **Max atlas frame size in luma samples**  **MaxAtlasSize** | **Max # of maps**  **LevelMapCount** | **Max # of attributes**  **MaxNumAttributeCount** | **Max # of attribute dimensions**  **MaxNumAttributeDims** |
| 1.0 | 30 000 000 | 1 000 000 | 2 228 224 | 2 | 1 | 3 |
| 2.0 | 60 000 000 | 2 000 000 | 2 228 224 | 2 | 3 | 3 |
| 3.0 |  |  | 8 912 896 | 2 | 1 | 3 |

Table A-8 specifies the general point cloud and VPS related limits for each level of each tier.

New variable/equation shall be less than or equal to MaxNumPointsPerSec, where MaxNumPointsPerSec is specified in Table A-8.

[Ed Note (all): check if there is a syntax element for that. If not, we need one or we need to come up with a way to derive this number.]

New variable/equation shall be less than or equal to MaxNumPointsPerFrame, where MaxNumPointsPerFrame is specified in Table A-8.

[Ed Note (all): Define variable in occupancy interpolation section that allows us to reference here. Should also include EOM and RAW points.]

asps\_frame\_width \* asps\_frame\_height shall be less than or equal to MaxAtlasSize., where MaxAtalasSize is specified in Table A-8.

vps\_map\_count\_minus1 shall be less than or equal to LevelMapCount – 1, where LevelMapCount is specified in Table A-8.

ai\_attribute\_count shall be less than or equal to MaxNumAttributeCount, where MaxNumAttributeCount is specified in Table A-8.

ai\_attribute\_dimension\_minus1 shall be less than or equal to MaxNumAttributeDims - 1, where MaxNumAttributeDims is specified in Table A-8.

Table A-9 — General atlas ASPS and tile related level limits

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Level** | **Max # Patches**  **MaxNumPatches** | **Max # RAW patches**  **MaxNumRawPatches** | **Max # RAW points**  **MaxNumRawPoints** | **Max # EOM patches**  **MaxNumEOMPatches** | **Max # EOM points**  **MaxNumEOMPoints** | **Max # CAB for atlas**  **MaxCAB** | **Max atlas bitrate**  **MaxAtlasBR** | **Max # of tile columns**  **MaxTileColumns** | **Max # of tile rows**  **MaxTileRows** |
| 1.0 | 256 | 32 | 50 000 | 32 | 50 000 |  | 1 500 000 | 11 | 10 |
| 2.0 | 256 | 64 | 50 000 | 32 | 50 000 |  | 2 000 000 | 11 | 10 |
| 3.0 | ? | ? | ? | ? | ? |  |  |  |  |

Table A-9 specifies the atlas ASPS and tile related level limits. The specified limits apply to all atlas tile groups forming the V-PCC bitstream.

MaxNumPatches is defined as the sum of all patches in an atlas frame.

globalPatchCounter shall be less than or equal to MaxNumPatches., where MaxNumPatches is specified in Table A-9.

[Ed Note (all): Refer to counters (global counter in atlas frame). Same for all below. If there’s no counter, define one. Same for below.]

globalRawPatchCounter shall be less than or equal to MaxNumRawPatches, where MaxNumRawPatches is specified in Table A-9.

globalRawPointsCounter shall be less than or equal to MaxNumRawPoints, where MaxNumRawPoints is specified in Table A-9.

globalEOMPatchCounter shall be less than or equal to MaxNumEOMPatches, where MaxNumEOMPatches is specified in Table A-9.

globalEOMPointsCounter shall be less than or equal to MaxNumEOMPoints, where MaxNumEOMPoints is specified in Table A-9.

[Ed Note (all): Do we want to combine EOM and RAW points in auxiliary points?]

MaxCAB, is defined as the number of atalas frame kept in the buffer.

[Ed Note (all): Check text with HEVC. Should it be defined in pictures or data units. Vlad will calculate something.]

MaxAtlasBR,…

afti\_num\_tile\_columns\_minus1 shall be less than or equal to MaxTileColumns – 1, where MaxTileColumns is specified in Table A-9.

afti\_num\_tile\_rows\_minus1 shall be less than or equal to MaxTileRows – 1, where MaxTileRows is specified in Table A-9.

[Ed Note (all): Shall we limit the patch distribution between tilegroups.]

Table A-10 — General video bitstream level limits

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Level** | **Max aggregate sample rate**  **MaxAggregateSr (samples/sec)** | **Max # Mbit/s per video stream**  **MaxBitratePerStream** | **Max aggregate # Mbit/s**  **MaxAggregateBitrate** | **Max # CPB per video stream**  **MaxCPBPerStream** |
| 1.0 | 133 693 440 | 50 | ? | 20 000 |
| 2.0 | 267 386 880 | 50 | ? | 20 000 |
| 3.X | 534 773 760 |  | 40 |  |

Table A-10 specifies the general video bitstream level limits.

[Ed. (All): Do we specify samples per second rate and bitrate per stream or as an aggregate. See M50998]

[Ed. (All): Is the bitrate defined according to the High Tier of HEVC? It seems too high]

[Ed. (SS): Regarding above, HEVC level 4.1 High (50Mbit/s)]

[Ed. (SS):Request for higher level requirements from industry.]

[Ed. (All): It was discussed whether we should consider having two level scenarios. A scenario controlled by individual bitstream decode capabilities and a second scenario controlled by aggregate bitstream capabilities. In the first, the limits are dictated for each sub-bitstream independently. For the second, the aggregate numbers are divided by the number of sub-bitstreams available. It might be desirable if both options are to be supported to either distinguish so through the level value or through an additional parameter in the profile/level syntax.]

1. Post-decoding conversion to nominal video formats
   1. Nominal video formats

Need a definition

* 1. Video format conversion process
     1. General process

[occupancy map] To convert the bitdepth of the occupancy map from the atlas defined by atlasId, to the required nominal bitdepth resolution, Section B.2.2 is called with inputs decOccFrame[ orderIdx ][ 0 ][ y ][ x ], and their associated bitdepth, decOccBitdepth[ orderIdx ], luma width, decOccWidth[ orderIdx ], and luma height, decOccHeight[ orderIdx ], the nominal bit depth defined by the syntax element oi\_occupancy\_nominal\_2d\_bitdepth\_minus1[ atlasId ], and the flag that indicates the MSB alignment, defined by the syntax element oi\_occupancy\_MSB\_align\_flag[ atlasId ].

[geometry map]

[attribute map]

* + 1. Conversion of a two dimensional array to a nominal bitdepth

Inputs to this process are:

* a two dimensional (inputWidth)x(inputHeight) array inputSamples, and its associated bitdepth, inputBitdepth, width, inputWidth, height, inputHeight
* the nominal bitdepth, nominalBitdepth, and a flag that indicates the MSB alignment, alignmentFlag

Ouput of this process is:

* a two dimensional (inputWidth)x(inputHeight) array outputSamples.

The variable bitdepthDifference is derived as follows:

bitdepthDifference = inputBitdepth – nominalBitdepth

Then the array outputSamples is generated in the following manner:

if(( bitdepthDifference > 0 ) && ( alignmentFlag == 1)){

for( y=0; y < inputHeight; y++ )  
 for( x=0; x < inputWidth; x++ )

outputSamples[ y ][ x ] = ( inputSamples[ y ][ x ] >> bitdepthDifference)

}

else if(( bitdepthDifference > 0 )){

for( y=0; y < inputHeight; y++ )  
 for( x=0; x < inputWidth; x++ )

outputSamples[ y ][ x ] = Min( inputSamples[ y ][ x ], (1 << ( nominalBitdepth + 1 ) ) – 1 )

}

else if( alignmentFlag == 1 ){

for( y=0; y < inputHeight; y++ )  
 for( x=0; x < inputWidth; x++ )

outputSamples[ y ][ x ] = ( inputSamples[ y ][ x ] << (-bitdepthDifference) )

}

else {

for( y=0; y < inputHeight; y++ )  
 for( x=0; x < inputWidth; x++ )

outputSamples[ y ][ x ] = inputSamples[ y ][ x ]

}

* + 1. Video format conversion process for resolution
    2. Video format conversion process for chroma upsampling
    3. Video format conversion process for frame rate
    4. Video format conversion process for attribute transfer

[Ed. (All): Copy aspects of section 9.2 and Annex A]

* 1. Geometry compositing process

Inputs to this process are the decoded geometry videos of an atlas with index atlasIdx after video format conversion as specified in section B.2 inpGeoFrame[inpMapIdx][ inpOrderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, inpGeoBitdepth, width, inpGeoWidth, and height, inpGeoHeight, where inpMapIdx corresponds to the value of the vuh\_map\_index of the input video and is in the range of 0 to vps\_map\_count\_minus1[ atlasIdx ]  − 1, inclusive, inpOrderIdx  is the display order index of all the decoded geometry frames, compIdx corresponds to the colour component index and is equal to 0, y is the column index in the decoded frame and is in the range of 0 to inpGeoHeight − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to tempGeoWidth − 1, inclusive.

Outputs of this process are the composited geometry video frames in display/output order, GeoFrame[ mapIdx ][ orderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, GeoBitdepth, luma width, GeoWidth, and luma height, GeoHeight, where mapIdx corresponds to the map index and is in the range of 0 to vps\_map\_count\_minus1[ atlasIdx ]  − 1, inclusive, orderIdx is the display order index of the decoded geometry frames, compIdx corresponds to the colour component index and is equal to 0, y is the row index in the decoded frame and is in the range of 0 to GeoHeight − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to GeoWidth − 1, inclusive.

The geometry video compositing process for the atlas,with index atlasIdx, is performed as follows:

First the output geometry bitdepth, width, and height are assigned as follows:

GeoBitdepth = inpGeoBitdepth

GeoWidth = inpGeoWidth

GeoHeight = inpGeoHeight

Afterwards, if vps\_map\_count\_minus1[ atlasIdx ] is equal to 0

for( y = 0; y < GeoHeight ; y++ )

for( x = 0; x < GeoWidth ; x++ )

GeoFrame[0][ orderIdx ][ 0 ][ y ][ x ]= inpGeoFrame[0][ orderIdx ][ y ][ x ]

Otherwise, (if vps\_map\_count\_minus1[ atlasIdx ] is not equal to 0), then

if( vps\_multiple\_map\_streams\_present\_flag[ atlasIdx ] is equal to 0 ) then:

for( i = 0; i <= vps\_map\_count\_minus1[ atlasIdx ]; i++ ){

mappedIdx = orderIdx \* ( vps\_map\_count\_minus1[ atlasIdx ] + 1 ) + i

for( y = 0; y < GeoHeight ; y++ )

for( x = 0; x < GeoWidth ; x++ )

GeoFrame[ i ][ orderIdx ][ 0 ][ y ][ x ]= inpGeoFrame[0][ mappedIdx ][ 0 ][ y ][ x ]

}

Otherwise:

geoMaxValue = (1 << GeoBitdepth) – 1

for( i = 0; i <= vps\_map\_count\_minus1[ atlasIdx ]; i++ ){

if( vps\_map\_absolute\_coding\_enabled\_flag[ atlasIdx ][ i ] == 0) {

j=  VPCCMapPredictorIndex( i )

for( y = 0; y < GeoHeight ; y++ )

for( x = 0; x < GeoWidth ; x++ )

GeoFrame[ i ][ orderIdx ][ 0 ][ y ][ x ] =

Clip3(0, geoMaxValue,

inpGeoFrame[ i ][ orderIdx ][ 0 ][ y ][ x ] +

GeoFrame[ j ][ orderIdx ][ 0 ][ y ][ x ] )

}

else {

for( y = 0; y < GeoHeight ; y++ )

for( x = 0; x < GeoWidth ; x++ )

GeoFrame[ i ][ orderIdx ][ 0 ][ y ][ x ]= inpGeoFrame[ i ][ orderIdx ][ 0 ][ y ][ x ]

}

}

* 1. Attribute compositing process

Inputs to this process are the decoded attribute videos of an atlas with index atlasIdx after video format conversion as specified in section B.2, inpAttrFrame[attrIdx][inpMapIdx][ inpOrderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, inpAttrBitdepth[attrIdx], width, inpAttrWidth, and height, inpAttrHeight, where attrIdx corresponds to the attribute index and is in the range of 0 to ai\_attribute\_count[ atlasIdx ]  − 1, inclusive, inpMapIdx corresponds to the value of the vuh\_map\_index of the input video and is in the range of 0 to vps\_map\_count\_minus1[ atlasIdx ]  − 1, inclusive, inpOrderIdx  is the display order index of all the decoded frames, compIdx corresponds to the attribute component index and is in the range of 0 to ai\_attribute\_dimension\_minus1[ attrIdx ], y is the column index in the attribute decoded frame and is in the range of 0 to inpAttrHeight − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to tempAttrWidth − 1, inclusive.

Outputs of this process are the composited attribute video frames in display/output order, AttrFrame[attrIdx][ mapIdx ][ orderIdx ][ compIdx ][ y ][ x ], and their associated bitdepth, AttrBitdepth[attrIdx], luma width, AttrWidth, and luma height, AttrHeight, where attrIdx corresponds to the attribute index and is in the range of 0 to ai\_attribute\_count[ atlasIdx ]  − 1, mapIdx corresponds to the map index and is in the range of 0 to vps\_map\_count\_minus1[ atlasIdx ]  − 1, inclusive, orderIdx is the display order index of the decoded attribute frames, compIdx corresponds to the attribute component index and is in the range of 0 to ai\_attribute\_dimension\_minus1[ attrIdx ], y is the row index in the decoded frame and is in the range of 0 to AttrHeight − 1, inclusive, and x is the column index in the decoded frame and is in the range of 0 to AttrWidth − 1, inclusive.

If ai\_attribute\_count[ atlasIdx ] is equal to 0, no attribute video frames are composited and no attribute information is associated with the final, reconstructed point cloud.

Otherwise, the following steps are performed for each attribute with index attrIdx for the atlas with index atlasIdx:

First the output attribute bitdepth, width, and height are assigned as follows:

AttrBitdepth[attrIdx] = inpAttrBitdepth[attrIdx]

AttrWidth = inpAttrWidth

AttrHeight = inpAttrHeight

The variable attributeDimension is then computed as follows:

attributeDimension = ai\_attribute\_dimension\_minus1[ atlasIdx ][ attrIdx ] + 1

Afterwards, if vps\_map\_count\_minus1[ atlasIdx ] is equal to 0

for( j = 0; j < attributeDimension; j++ ){

for( y = 0; y < AttrHeight ; y++ )

for( x = 0; x < AttrWidth ; x++ )

AttrFrame[attrIdx][0][ orderIdx ][ 0 ][j][ y ][ x ]=

 inpAttrFrame[attrIdx][0][ orderIdx ][j][ y ][ x ]

Otherwise, (if vps\_map\_count\_minus1[ atlasIdx ] is not equal to 0), then

if (vps\_multiple\_map\_streams\_present\_flag[ atlasIdx ] is equal to 0) then:

for( i = 0; i <= vps\_map\_count\_minus1[ atlasIdx ]; i++ ){

mappedIdx = orderIdx \* ( vps\_map\_count\_minus1[ atlasIdx ] + 1 ) + i

for( j = 0; j < attributeDimension; j++ ){

for( y = 0; y < AttrHeight ; y++ )

for( x = 0; x < AttrWidth ; x++ )

AttrFrame[ attrIdx ][ i ][ orderIdx ][ j ][ y ][ x ] =

inpAttrFrame[ attrIdx ][0][ mappedIdx ][ j ][ y ][ x ]

}

}

Otherwise:

attrMaxValue = (1 << AttrBitdepth[attrIdx]) – 1

attrOffset = (1 << (AttrBitdepth[attrIdx] – 1 )

for( i = 0; i <= vps\_map\_count\_minus1[ atlasIdx ]; i++ ){

if( AttributeMapAbsoluteCodingEnabledFlag[ atlasIdx ][ attrIdx ][ i ] == 0) {

j=  VPCCMapPredictorIndex( i )

for( k = 0; k < attributeDimension; k++ ){

for( y = 0; y < AttrHeight ; y++ )

for( x = 0; x < AttrWidth ; x++ )

AttrFrame[attrIdx][ i ][ orderIdx ][ k ][ y ][ x ] =

Clip3(0, attrMaxValue, – attrOffset

inpAttrFrame[attrIdx][ i ][ orderIdx ][ k ][ y ][ x ] +

AttrFrame[attrIdx][ j ][ orderIdx ][ k ][ y ][ x ] )

}

}

else {

for( k = 0; k < attributeDimension; k++ ){

for( y = 0; y < AttrHeight ; y++ )

for( x = 0; x < AttrWidth ; x++ )

AttrFrame[attrIdx][ i ][ orderIdx ][ k ][ y ][ x ]=

inpAttrFrame[attrIdx][ i ][ orderIdx ][ k ][ y ][ x ]

}

}

* 1. Attribute dimension packing process for an attribute with attrIdx

Input to this process is the atlas index, atlasIdx, attribute index, attrIdx, a variable, orderIdx, corresponding to the display order of a frame in a decoded video, a variable, numberOfPartitions, corresponding to a number of dimension partitions in which an attribute has been partitioned, a series of frames inpFrame[ partIdx ][ orderIdx ][ inpCompIdx ][ y ][ x ], and their associated width, inpWidth, and height, inpHeight, information, where partIdx is in the range of 0 to numberOfPartitions – 1, inclusive, and inpCompIdx is in the range of 0 to ai\_attribute\_partition\_channels\_minus1[ atlasIdx ][ attrIdx ][ partIdx ], inclusive.

Output to this process is a frame outFrame[ orderIdx ][ compIdx ][ y ][ x ] where compIdx is in the range of 0 to ai\_attribute\_dimension\_minus1[ atlasIdx ][ attrIdx ] – 1.

outComp = 0

for( m = 0; m < numberOfPartitions; m++, outComp++ ) {

for(k = 0; k < ai\_attribute\_partition\_channels\_minus1[ atlasIdx ][ attrIdx ][ partIdx ] + 1; k++ ) {

for( j = 0; j < inpHeight; j++ )

for( i = 0; i < inpWidth; i++ ) {

outFrame[ orderIdx ][ outComp ][ j ][ i ] = inpFrame[ m ][ orderIdx ][ k ][ j ][ i ]

}

}

}

1. V-PCC Sample stream format

(This annex forms an integral part of this Recommendation | International Standard.)

* 1. General

This annex specifies syntax and semantics of a sample stream format specified for use by applications that deliver some or all of the V-PCC unit stream as an ordered stream of bytes or bits within which the locations of V-PCC unit boundaries need to be identifiable from patterns in the data, such as Rec. ITU-T H.222.0 | ISO/IEC 13818-1 systems or Recommendation ITU‑T H.320 systems. For bit-oriented delivery, the bit order for the sample stream format is specified to start with the MSB of the first byte, proceed to the LSB of the first byte, followed by the MSB of the second byte, etc.

The sample stream format starts with a sample stream header and consists of a sequence of sample stream V-PCC unit syntax structures. Each sample stream V-PCC unit syntax structure contains one element, named ssvu\_vpcc\_unit\_size,that specifies a size, followed by one vpcc\_unit( ssvu\_vpcc\_unit\_size) syntax structure. This vpcc\_unit is essentially of size ssvu\_vpcc\_unit\_size. The number of bits used to encode the element ssvu\_vpcc\_unit\_size is specified in the sample stream header.

* 1. Sample stream V-PCC unit syntax and semantics
     1. Sample stream V-PCC header syntax

|  |  |
| --- | --- |
| sample\_stream\_vpcc\_header() { | **Descriptor** |
| **ssvh\_unit\_size\_precision\_bytes\_minus1** | u(3) |
| **ssvh\_reserved\_zero\_5bits** | u(5) |
| } |  |

* + 1. Sample stream V-PCC unit syntax

|  |  |
| --- | --- |
| sample\_stream\_vpcc\_unit() { | **Descriptor** |
| **ssvu\_vpcc\_unit\_size** | u(v) |
| vpcc\_unit(ssvu\_vpcc\_unit\_size) |  |
| } |  |

* + 1. Sample stream V-PCC header semantics

The sample stream V-PCC header shall always be at the beginning of the sample stream V-PCC stream.

**ssvh\_unit\_size\_precision\_bytes\_minus1** plus 1 specifies the precision, in bytes, of the ssvu\_vpcc\_unit\_size element in all sample stream V-PCC units. ssvh\_unit\_size\_precision\_bytes\_minus1 shall be in the range of 0 to 7.

**ssvh\_reserved\_zero\_5bits** shall be equal to 0 in bitstreams conforming to this version of this document. Other values for sshv\_reserved\_zero\_5bits are reserved for future use by ISO/IEC. Decoders shall ignore the value of sshv\_reserved\_zero\_5bits.

* + 1. Sample stream V-PCC unit semantics

The order of sample stream V‑PCC units in the sample stream shall follow the decoding order of the V‑PCC units contained in the sample stream V‑PCC units (see clause xxx). The content of each sample stream V‑PCC unit is associated with the same access unit as the V‑PCC unit contained in the sample stream V‑PCC unit (see clause xxxxx).

**ssvu\_vpcc\_unit\_size** specifies the size, in bytes, of the subsequent vpcc\_unit. The number of bits used to represent ssvu\_vpcc\_unit\_size is equal to (ssvh\_unit\_size\_precision\_bytes\_minus1 + 1) \* 8.

1. NAL Sample stream format

(This annex forms an integral part of this Recommendation | International Standard.)

* 1. General

This annex specifies syntax and semantics of a sample stream format specified for use by applications that deliver some or all of the atlas NAL unit stream as an ordered stream of bytes or bits within which the locations of atlas NAL unit boundaries need to be identifiable from patterns in the data, such as Rec. ITU-T H.222.0 | ISO/IEC 13818-1 systems or Recommendation ITU‑T H.320 systems. For bit-oriented delivery, the bit order for the sample stream format is specified to start with the MSB of the first byte, proceed to the LSB of the first byte, followed by the MSB of the second byte, etc.

The sample stream format starts with a sample stream header and consists of a sequence of sample stream NAL unit syntax structures. Each sample stream NAL unit syntax structure contains one element, named ssnu\_nal\_unit\_size,that specifies a size, followed by one nal\_unit( ssnu\_nal\_unit\_size) syntax structure. This nal\_unit is essentially of size ssnu\_nal\_unit\_size. The number of bits used to encode the element ssnu\_nal\_unit\_size is specified in the sample stream header.

* 1. Sample stream NAL unit syntax and semantics
     1. Sample stream NAL header syntax

|  |  |
| --- | --- |
| sample\_stream\_nal\_header() { | **Descriptor** |
| **ssnh\_unit\_size\_precision\_bytes\_minus1** | u(3) |
| **ssnh\_reserved\_zero\_5bits** | u(5) |
| } |  |

* + 1. Sample stream NAL unit syntax

|  |  |
| --- | --- |
| sample\_stream\_nal\_unit() { | **Descriptor** |
| **ssnu\_nal\_unit\_size** | u(v) |
| nal\_unit(ssnu\_nal\_unit\_size) |  |
| } |  |

* + 1. Sample stream NAL header semantics

The sample stream NAL header shall always be at the beginning of the NAL stream.

**ssnh\_unit\_size\_precision\_bytes\_minus1** plus 1 specifies the precision, in bytes, of the ssnu\_nal\_unit\_size element in all sample stream NAL units. ssnh\_unit\_size\_precision\_bytes\_minus1 shall be in the range of 0 to 7.

**ssnh\_reserved\_zero\_5bits** shall be equal to 0 in bitstreams conforming to this version of this document. Other values for ssnh\_reserved\_zero\_5bits are reserved for future use by ISO/IEC. Decoders shall ignore the value of ssnh\_reserved\_zero\_5bits.

* + 1. Sample stream NAL unit semantics

The order of sample stream NAL units in the sample stream shall follow the decoding order of the NAL units contained in the sample stream NAL units (see clause xxx). The content of each sample stream NAL unit is associated with the same access unit as the NAL unit contained in the sample stream NAL unit (see clause xxxxx).

**ssnu\_nal\_unit\_size** specifies the size, in bytes, of the subsequent NAL\_unit. The number of bits used to represent ssnu\_nal\_unit\_size is equal to (ssnh\_unit\_size\_precision\_bytes\_minus1 + 1) \* 8.

1. Atlas Hypothetical Reference Decoder

(This annex forms an integral part of this Recommendation | International Standard.)

* 1. General

This annex specifies the hypothetical reference decoder (HRD) and its use to check atlas bitstream and decoder conformance.

Two types of bitstreams or bitstream subsets are subject to HRD conformance checking for this document. The first type, called a Type I bitstream, is an ACL NAL unit stream containing only the ACL NAL units and NAL units with nal\_unit\_type equal to NAL\_FD (filler data NAL units) for all access units in the bitstream. The second type, called a Type II bitstream, contains, in addition to the ACL NAL units and filler data NAL units for all access units in the atlas bitstream, additional non-ACL NAL unitsother than filler data units.

Figure E-1 shows the types of bitstream conformance points checked by the atlas sub-bitstream HRD.



Figure E-1 – Structure of Atlas NAL unit streams for HRD conformance checks

The syntax elements of non-ACL NAL units (or their default values for some of the syntax elements), required for the HRD, are specified in the semantic clauses of clause 7 and Annex F.

Two types of HRD parameter sets (NAL HRD parameters and non-ACL HRD parameters) are used. The HRD parameter sets are signalled through the hrd\_parameters ( ) syntax structure, which is part of the VUI syntax structure.

Multiple tests may be needed for checking the conformance of an atlas sub-bitstream, which is referred to as the bitstream under test. For each test, the following steps apply in the order listed:

1. An operation point under test, denoted as TargetOp, is set by selecting a target highest TemporalId value OpTid.
2. The hrd\_parameters( ) syntax structure in the VUI of the active ASPS (or provided through some external means not specified in this document) that applies to TargetOp that is selected. Within the selected hrd\_parameters( ) syntax structure, if AtlasBitstreamToDecode is a Type I bitstream the variable NalHrdModeFlag is set equal to 0; otherwise (AtlasBitstreamToDecode is a Type II bitstream). When AtlasBitstreamToDecode is a Type II bitstream and NalHrdModeFlag is equal to 0, all non-ACL NAL units except filler data NAL units from the NAL unit stream (as specified in Annex C), when present, are discarded from AtlasBitstreamToDecode and the remaining atlas bitstream is assigned to AtlasBitstreamToDecode.
3. A coded atlas access unit associated with a buffering period SEI message (present in AtlasBitstreamToDecode or available through external means not specified in this document) applicable to TargetOp is selected as the HRD initialization point and referred to as atlas access unit 0.
4. For each atlas access unit in AtlasBitstreamToDecode starting from atlas access unit 0, the buffering period SEI message (present in AtlasBitstreamToDecode or available through external means not specified in this document) that is associated with the atlas access unit and applies to TargetOp is selected, the atlas frame timing SEI message (present in AtlasBitstreamToDecode or available through external means not specified in this document) that is associated with the atlas access unit and applies to TargetOp is selected.
5. A value of SchedSelIdx is selected. The selected SchedSelIdx shall be in the range of 0 to hrd\_cab\_cnt\_minus1 inclusive, where hrd\_cab\_cnt\_minus1 is found in the hrd\_parameters( ) syntax structure as selected above.
6. When the coded atlas frame in access unit 0 has nal\_unit\_type equal to NAL\_CRA/GCRA or NAL\_BLA\_W\_LP/NAL\_GBLA\_W\_LP, and bp\_irap\_cab\_params\_present\_flag in the selected buffering period SEI message is equal to 1, either of the following applies for selection of the initial CAB removal delay and delay offset:
   * If NalHrdModeFlag is equal to 1, the default initial CAB removal delay and the delay offset represented by bp\_nal\_initial\_cab\_removal\_delay[ SchedSelIdx ] and bp\_nal\_initial\_cab\_removal\_offset[ SchedSelIdx ], respectively, in the selected buffering period SEI message are selected. Otherwise, the default initial CAB removal delay and delay offset represented by bp\_acl\_initial\_cab\_removal\_delay[ SchedSelIdx ] and bp\_acl\_initial\_cab\_removal\_offset[ SchedSelIdx ], respectively, in the selected buffering period SEI message are selected. The variable DefaultInitCabParamsFlag is set equal to 1.
   * If NalHrdModeFlag is equal to 1, the alternative initial CAB removal delay and delay offset represented by bp\_nal\_initial\_alt\_cab\_removal\_delay[ SchedSelIdx ] and bp\_nal\_initial\_alt\_cab\_removal\_offset[ SchedSelIdx ], respectively, in the selected buffering period SEI message are selected. Otherwise, the alternative initial CAB removal delay and delay offset represented by bp\_acl\_initial\_alt\_cab\_removal\_delay[ SchedSelIdx ] and bp\_acl\_initial\_alt\_cab\_removal\_offset[ SchedSelIdx ], respectively, in the selected buffering period SEI message are selected. The variable DefaultInitCafbParamsFlag is set equal to 0, and the RASL access units associated with access unit 0 are discarded from AtlasBitstreamToDecode and the remaining bitstream is assigned to AtlasBitstreamToDecode.

Each conformance test consists of a combination of one option in each of the above steps. When there is more than one option for a step, for any conformance test only one option is chosen. All possible combinations of all the steps form the entire set of conformance tests. For each operation point under test, the number of bitstream conformance tests to be performed is equal to n0 \* n1 \* ( n2 \* 2 + n3 ) \* n4, where the values of n0, n1, n2, n3 and n4 are specified as follows:

* n0 is derived as follows:
  + If AtlasBitstreamToDecode is a Type I bitstream, n0 is equal to 1.
  + Otherwise (AtlasBitstreamToDecode is a Type II bitstream), n0 is equal to 2.
* n1 is equal to hrd\_cab\_cnt\_minus1 + 1.
* n2 is the number of access units in AtlasBitstreamToDecode that each is associated with a buffering period SEI message applicable to TargetOp and for each of which both of the following conditions are true:
  + nal\_unit\_type is equal to NAL\_GCRA/GBLA/ NAL\_CRA/BLAfor the ACL NAL units.
  + The associated buffering period SEI message applicable to TargetOp has bp\_irap\_cab\_params\_present\_flag equal to 1.
* n3 is the number of access units in AtlasBitstreamToDecode that each is associated with a buffering period SEI message applicable to TargetOp and for each of which one or both of the following conditions are true:
  + nal\_unit\_type is not equal to NAL\_GCRA, NAL\_GBLA, NAL\_CRA, or NAL\_BLA for the ACL NAL units.
  + The associated buffering period SEI message applicable to TargetOp has bp\_irap\_cab\_params\_present\_flag equal to 0.
* n4 is equal to 2.

When AtlasBitstreamToDecode is a Type II patch bitstream, the following applies:

* If the hrd\_parameters( ) syntax structure that immediately follows the condition "if( bp\_acl\_hrd\_parameters\_present\_flag )" is selected, the test is conducted at the Type I conformance point shown in Figure E-1, and only ACL and filler data NAL units are counted for the input bit rate and CAB storage.
* Otherwise (the hrd\_parameters( ) syntax structure that immediately follows the condition "if( nal\_hrd\_parameters\_present\_flag )" is selected), the test is conducted at the Type II conformance point shown in Figure E-1, and all bytes of the Type II bitstream, which is a NAL unit stream, are counted for the input bit rate and CAB storage.

NOTE 1 – NAL HRD parameters established by a value of SchedSelIdx for the Type II conformance point shown in Figure D-1 are sufficient to also establish ACL HRD conformance for the Type I conformance point shown in Figure D-1 for the same values of InitCabRemovalDelay[ SchedSelIdx ], BitRate[ SchedSelIdx ] and CabSize[ SchedSelIdx ] for the variable bit rate (VBR) case (hrd\_cbr\_flag[ SchedSelIdx ] equal to 0). This is because the data flow into the Type I conformance point is a subset of the data flow into the Type II conformance point and because, for the VBR case, the CAB is allowed to become empty and stay empty until the time a next atlas is scheduled to begin to arrive. For example, when decoding a CAS conforming to one or more of the profiles specified in Annex A using the decoding process specified in clause 8.4.

All ASPSs, and AFPSs referred to in the ACL NAL units and the corresponding buffering period and atlas timing information SEI messages shall be conveyed to the HRD, in a timely manner, either in the bitstream (by non-ACL NAL units), or by other means not specified in this document.

In Annex E and Annex F, the specification for "presence" of non-ACL NAL units that contain ASPSs, and AFPSs~~,~~ buffering period SEI messages, atlas frame timing SEI messages is also satisfied when those NAL units (or just some of them) are conveyed to atlas frame decoders (or to the HRD) by other means not specified in this document. For the purpose of counting bits, only the appropriate bits that are actually present in the bitstream are counted.

NOTE 2 – As an example, synchronization of such a non-ACL NAL unit, conveyed by means other than presence in the bitstream, with the NAL units that are present in the bitstream, can be achieved by indicating two points in the bitstream, between which the non‑ACL NAL unit would have been present in the bitstream, had the atlas frame encoder decided to convey it in the bitstream.

When the content of such a non-ACL NAL unit is conveyed for the application by some means other than presence within the bitstream, the representation of the content of the non-ACL NAL unit is not required to use the same syntax as specified in this document.

NOTE 3 – When HRD information is contained within the bitstream, it is possible to verify the conformance of a bitstream to the requirements of this clause based solely on information contained in the bitstream. When the HRD information is not present in the bitstream, as is the case for all "stand-alone" Type I bitstreams, conformance can only be verified when the HRD data are supplied by some other means not specified in this document.

The HRD contains a coded atlas buffer (CAB), an instantaneous decoding process, a decoded atlas buffer (DAB), and output cropping as shown in Figure E-2.



Figure E-2 – Atlas Frame HRD buffer model

For each atlas bitstream conformance test, the CAB size (number of bits) is CabSize[ SchedSelIdx ] as specified in clause G.3.3, where SchedSelIdx and the HRD parameters are specified above in this clause. The DAB size (number of atlas frames storage buffers) is asps\_max\_dec\_atlas\_frame\_buffering\_minus1 + 1.

The following is specified for expressing the constraints in this annex:

– Each access unit is referred to as access unit n, where the number n identifies the particular access unit. Access unit 0 is selected per step 3 above. The value of n is incremented by 1 for each subsequent access unit in decoding order.

– Atlas frame n refers to the coded atlas frame or the decoded atlas frame of access unit n.

The HRD operates as follows:

– The HRD is initialized at decoding unit 0, with both the CAB and the DAB being set to be empty (the DAB fullness is set equal to 0).

NOTE 4 – After initialization, the HRD is not initialized again by subsequent buffering period SEI messages.

– Data associated with decoding units that flow into the CAB according to a specified arrival schedule are delivered by the hypothetical atlas stream scheduler (HSS).

– The data associated with each decoding unit are removed and decoded instantaneously by the instantaneous decoding process at the CAB removal time of the decoding unit.

– Each decoded atlas frame is placed in the DAB.

– A decoded atlas frame is removed from the DAB when it becomes no longer needed for inter prediction reference and no longer needed for output.

For each bitstream conformance test, the operation of the CAB is specified in clause E.2, the instantaneous decoder operation is specified in clauses 2 through 10, the operation of the DAB is specified in clause E.3

HSS and HRD information concerning the number of enumerated delivery schedules and their associated bit rates and buffer sizes is specified in clauses G.2.2 and G.3.2. The HRD is initialized as specified by the buffering period SEI message specified in clauses  and  . The removal timing of decoding units from the CAB and output timing of decoded atlases from the DAB is specified using information in atlas frame timing SEI messages (specified in clauses E.2.3 and E.3.3). All timing information relating to a specific decoding unit shall arrive prior to the CAB removal time of the decoding unit.

The requirements for atlas sub-bitstream conformance are specified in clause E.4 and the HRD is used to check conformance of atlas sub-bitstreams as specified above in this clause and to check conformance of decoders as specified in clause E.5.

NOTE 5 – While conformance is guaranteed under the assumption that all “atlas frame”-rates and clocks used to generate the atlas sub-bitstream match exactly the values signalled in the atlas sub-bitstream, in a real system each of these may vary from the signalled or specified value.

All the arithmetic in this annex is performed with real values, so that no rounding errors can propagate. For example, the number of bits in a CAB just prior to or after removal of a decoding unit is not necessarily an integer.

The variable ClockTick is derived as follows and is called a clock tick:

ClockTick = vui\_num\_units\_in\_tick ÷ vui\_time\_scale (E‑1)

* 1. Operation of coded atlas frame buffer
     1. General

The specifications in this clause apply independently to each set of coded atlas frame buffer (CAB) parameters that is present and to both the Type I and Type II conformance points shown in Figure E-1 and the set of CAB parameters is selected as specified in clause E.1.

* + 1. Timing of decoding unit arrival

The decoding unit is considered as an access unit, for derivation of the initial and final CAB arrival times for access unit n.

The variables InitCabRemovalDelay[ SchedSelIdx ] and InitCabRemovalDelayOffset[ SchedSelIdx ] are derived as follows:

* If one or more of the following conditions are true, InitCabRemovalDelay[ SchedSelIdx ] and InitCabRemovalDelayOffset[ SchedSelIdx ] are set equal to the values of the buffering period SEI message syntax elements bp\_nal\_initial\_alt\_cab\_removal\_delay[ SchedSelIdx ] and bp\_nal\_initial\_alt\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 1 or bp\_acl\_initial\_alt\_cab\_removal\_delay[ SchedSelIdx ] and bp\_acl\_initial\_alt\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 0, where the buffering period SEI message syntax elements are selected as specified in clause E.1:
  + Access unit 0 is a NAL\_CRA/GCRA/NAL\_BLA/GBLA access unit, and the value of bp\_irap\_cab\_params\_present\_flag of the buffering period SEI message is equal to 1 and one or more of the following conditions are true:
    - DefaultInitCabParamsFlag is equal to 0.
* Otherwise, InitCabRemovalDelay[ SchedSelIdx ] and InitCabRemovalDelayOffset[ SchedSelIdx ] are set equal to the values of the buffering period SEI message syntax elements bp\_nal\_initial\_cab\_removal\_delay[ SchedSelIdx ] and bp\_nal\_initial\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 1, or bp\_acl\_initial\_cab\_removal\_delay[ SchedSelIdx ] and bp\_acl\_initial\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 0, where the buffering period SEI message syntax elements are selected as specified in clause E.1.

The time at which the first bit of decoding unit m begins to enter the CAB is referred to as the initial arrival time initArrivalTime[ m ].

The initial arrival time of decoding unit m is derived as follows:

* If the decoding unit is decoding unit 0 (i.e., when m is equal to 0), initArrivalTime[ 0 ] is set equal to 0.
* Otherwise (the decoding unit is decoding unit m with m > 0), the following applies:
  + If hrd\_cbr\_flag[ SchedSelIdx ] is equal to 1, the initial arrival time for decoding unit m is equal to the final arrival time (which is derived below) of decoding unit m − 1, i.e.,

initArrivalTime[ m ] = AuFinalArrivalTime[ m − 1 ] (E‑2)

* Otherwise (hrd\_cbr\_flag[ SchedSelIdx ] is equal to 0), the initial arrival time for decoding unit m is derived as follows:
* initArrivalTime[ m ] = Max( AuFinalArrivalTime[ m − 1 ], initArrivalEarliestTime[ m ] ) (E‑3)

where initArrivalEarliestTime[ m ] is derived as follows:

* + The variable tmpNominalRemovalTime is derived as follows:
* tmpNominalRemovalTime = AuNominalRemovalTime[ m ] (E‑4)

where AuNominalRemovalTime[ m ] is the nominal CAB removal time of access unit m respectively, as specified in clause E.2.3.

* If decoding unit m is not the first decoding unit of a subsequent buffering period, initArrivalEarliestTime[ m ] is derived as follows:
* initArrivalEarliestTime[ m ] = tmpNominalRemovalTime −   
   ( InitCabRemovalDelay[ SchedSelIdx ] +   
   InitCabRemovalDelayOffset[ SchedSelIdx ] ) ÷ 90000 (E‑5)
* Otherwise (decoding unit m is the first decoding unit of a subsequent buffering period), initArrivalEarliestTime[ m ] is derived as follows:
* initArrivalEarliestTime[ m ] = tmpNominalRemovalTime −  
   ( InitCabRemovalDelay[ SchedSelIdx ] ÷ 90000 ) (E‑6)

The final arrival time for decoding unit m is derived as follows:

AuFinalArrivalTime[ m ] = initArrivalTime[ m ] + sizeInbits[ m ] ÷   
 BitRate[ SchedSelIdx ] (E‑7)

where sizeInbits[ m ] is the size in bits of decoding unit m, counting the bits of the ACL NAL units and the filler data NAL units for the Type I conformance point or all bits of the Type II bitstream for the Type II conformance point, where the Type I and Type II conformance points are as shown in Figure E-1.

The values of SchedSelIdx, BitRate[ SchedSelIdx ] and CabSize[ SchedSelIdx ] are constrained as follows:

* If the content of the selected hrd\_parameters( ) syntax structures for the access unit containing decoding unit m and the previous access unit differ, the HSS selects a value SchedSelIdx1 of SchedSelIdx from among the values of SchedSelIdx provided in the selected hrd\_parameters( ) syntax structures for the access unit containing decoding unit m that results in a BitRate[ SchedSelIdx1 ] or CabSize[ SchedSelIdx1 ] for the access unit containing decoding unit m. The value of BitRate[ SchedSelIdx1 ] or CabSize[ SchedSelIdx1 ] may differ from the value of BitRate[ SchedSelIdx0 ] or CabSize[ SchedSelIdx0 ] for the value SchedSelIdx0 of SchedSelIdx that was in use for the previous access unit.
* Otherwise, the HSS continues to operate with the previous values of SchedSelIdx, BitRate[ SchedSelIdx ] and CabSize[ SchedSelIdx ].

When the HSS selects values of BitRate[ SchedSelIdx ] or CabSize[ SchedSelIdx ] that differ from those of the previous access unit, the following applies:

* The variable BitRate[ SchedSelIdx ] comes into effect at the initial CAB arrival time of the current access unit.
* The variable CabSize[ SchedSelIdx ] comes into effect as follows:
  + If the new value of CabSize[ SchedSelIdx ] is greater than the old CAB size, it comes into effect at the initial CAB arrival time of the current access unit.
  + Otherwise, the new value of CabSize[ SchedSelIdx ] comes into effect at the CAB removal time of the current access unit.
    1. Timing of decoding unit removal and decoding of decoding unit

The variables InitCabRemovalDelay[ SchedSelIdx ], InitCabRemovalDelayOffset[ SchedSelIdx ], CabDelayOffset and DabDelayOffset are derived as follows:

* If one or more of the following conditions are true, CabDelayOffset is set equal to the value of the buffering period SEI message syntax element bp\_cab\_delay\_offset, DabDelayOffset is set equal to the value of the buffering period SEI message syntax element bp\_dab\_delay\_offset, and InitCabRemovalDelay[ SchedSelIdx ] and InitCabRemovalDelayOffset[ SchedSelIdx ] are set equal to the values of the buffering period SEI message syntax elements bp\_nal\_initial\_alt\_cab\_removal\_delay[ SchedSelIdx ] and bp\_nal\_initial\_alt\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 1, or bp\_acl\_initial\_alt\_cab\_removal\_delay[ SchedSelIdx ] and bp\_acl\_initial\_alt\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 0, where the buffering period SEI message containing the syntax elements is selected as specified in clause E.1.
  + Access unit 0 is a BLA access unit for which the coded atlas has nal\_unit\_type equal to NAL\_BLA\_W\_RADL or NAL\_BLA\_N\_LP, and the value of bp\_irap\_cab\_params\_present\_flag of the buffering period SEI message is equal to 1.
  + Access unit 0 is a BLA access unit for which the coded atlas has nal\_unit\_type equal to NAL\_BLA\_W\_LP or is a NAL\_CRA access unit, and the value of bp\_irap\_cab\_params\_present\_flag of the buffering period SEI message is equal to 1, and one or more of the following conditions are true:
    - DefaultInitCabParamsFlag is equal to 0.
* Otherwise, InitCabRemovalDelay[ SchedSelIdx ] and InitCabRemovalDelayOffset[ SchedSelIdx ] are set equal to the values of the buffering period SEI message syntax elements bp\_nal\_initial\_cab\_removal\_delay[ SchedSelIdx ] and bp\_nal\_initial\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 1, or bp\_acl\_initial\_cab\_removal\_delay[ SchedSelIdx ] and bp\_acl\_initial\_cab\_removal\_offset[ SchedSelIdx ], respectively, when NalHrdModeFlag is equal to 0, where the buffering period SEI message containing the syntax elements is selected as specified in clause E.1, CabDelayOffset and DpbDelayOffset are both set equal to 0.

The nominal removal time of the access unit n from the CAB is specified as follows:

* If access unit n is the access unit with n equal to 0 (the access unit that initializes the HRD), the nominal removal time of the access unit from the CAB is specified by:

AuNominalRemovalTime[ 0 ] = InitCabRemovalDelay[ SchedSelIdx ] ÷ 90000 (E‑8)

* Otherwise, the following applies:
* When access unit n is the first access unit of a buffering period that does not initialize the HRD, the following applies:

The nominal removal time of the access unit n from the CAB is specified by:

if( !concatenationFlag ) {  
 baseTime = AuNominalRemovalTime[ firstAtlasInPrevBuffPeriod ]  
 tmpCabRemovalDelay = AuCabRemovalDelayVal  
 }  
 else {  
 baseTime = AuNominalRemovalTime[ prevNonDiscardablePic ]  
 tmpCabRemovalDelay =Max( ( auCabRemovalDelayDeltaMinus1 + 1 ), (E‑9)  
 Ceil( ( InitCabRemovalDelay[ SchedSelIdx ] ÷ 90000 +   
 AuFinalArrivalTime[ n − 1 ] − AuNominalRemovalTime[ n − 1 ] ) ÷ ClockTick )  
 }  
 AuNominalRemovalTime( n ) = baseTime + ClockTick \* ( tmpCabRemovalDelay −   
 CabDelayOffset )

where AuNominalRemovalTime[ firstAtlasInPrevBuffPeriod  ] is the nominal removal time of the first access unit of the previous buffering period, AuNominalRemovalTime[ prevNonDiscardableAtlas ] is the nominal removal time of the preceding atlas in decoding order with TemporalId equal to 0 that is not a RASL, RADL or sub-layer non-reference atlas frame, AuCabRemovalDelayVal is the value of AuCabRemovalDelayVal derived according to aft\_cab\_removal\_delay\_minus1 in the atlas timing SEI message, selected as specified in clause E.1, associated with access unit n, and concatenationFlag and auCabRemovalDelayDeltaMinus1 are the values of the syntax elements bp\_concatenation\_flag and bp\_atlas\_cab\_removal\_delay\_delta\_minus1, respectively, in the buffering period SEI message, selected as specified in clause E.1, associated with access unit n.

After the derivation of the nominal CAB removal time and before the derivation of the DAB output time of access unit n, the values of CabDelayOffset and DabDelayOffset are updated as follows:

* If one or more of the following conditions are true, CabDelayOffset is set equal to the value of the buffering period SEI message syntax element bp\_cab\_delay\_offset, and DabDelayOffset is set equal to the value of the buffering period SEI message syntax element bp\_dab\_delay\_offset, where the buffering period SEI message containing the syntax elements is selected as specified in clause E.1:
* Access unit n is a BLA access unit for which the coded atlas has nal\_unit\_type equal to BLA\_W\_RADL or BLA\_N\_LP, and the value of bp\_irap\_cab\_params\_present\_flag of the buffering period SEI message is equal to 1.
* Access unit n is a BLA access unit for which the coded atlas has nal\_unit\_type equal to BLA\_W\_LP or is a CRA access unit, and the value of bp\_irap\_cab\_params\_present\_flag of the buffering period SEI message is equal to 1, and UseAltCabParamsFlag for access unit n is equal to 1.
* Otherwise, CabDelayOffset and DpbDelayOffset are both set equal to 0.
* When access unit n is not the first access unit of a buffering period, the nominal removal time of the access unit n from the CAB is specified by:
* AuNominalRemovalTime[ n ] = AuNominalRemovalTime [ firstAtlasInCurrBuffPeriod ] +   
   ClockTick \* ( AuCabRemovalDelayVal − CabDelayOffset ) (E‑10)

where AuNominalRemovalTime[ firstAtlasInCurrBuffPeriod ] is the nominal removal time of the first access unit of the current buffering period, and AuCabRemovalDelayVal is the value of AuCabRemovalDelayVal derived according to aft\_cab\_removal\_delay\_minus1 in the atlas timing SEI message, selected as specified in clause E.1, associated with access unit n.

CAB removal time of access unit n, the access unit is instantaneously decoded.

* 1. Operation of the decoded atlas frame buffer
     1. General

The specifications in this clause apply independently to each set of decoded atlas frame buffer (DAB) parameters selected as specified in clause E.1.

The decoded atlas frame buffer contains atlas frame storage buffers. Each of the atlas frame storage buffers may contain a decoded atlas frame that is marked as "used for reference" or is held for future output. The processes specified in clauses E.3.2 and E.3.3 are sequentially applied as specified below.

* + 1. Removal of atlas frames from the DAB before decoding of the current atlas frame

The removal of atlas frames from the DAB before decoding of the current atlas frame (but after parsing the tile group header of the first tile group of the current atlas frame) happens instantaneously at the CAB removal time of the first decoding unit of access unit n (containing the current atlas frame) and proceeds as follows:

* The decoding process for reference atlas list (RAL) as specified in clause  8.4.3.2 is invoked.
* When the current atlas is an IRAP atlas frame with NoRaslOutputFlag equal to 1 that is not atlas frame 0, the following ordered steps are applied:

1. The variable NoOutputOfPriorAtlasFramesFlag is derived for the decoder under test as follows:
   * If the current atlas frame is a CRA atlas frame, NoOutputOfPriorAtlasFramesFlag is set equal to 1.
   * Otherwise, NoOutputOfPriorAtlasFramesFlag is set equal to 0.
2. The value of NoOutputOfPriorAtlasFramesFlag derived for the decoder under test is applied for the HRD, such that when the value of NoOutputOfPriorAtlasFramesFlag is equal to 1, all atlas storage buffers in the DAB are emptied without output of the atlas frames they contain, and the DAB fullness is set equal to 0.
   * When both of the following conditions are true for any atlass k in the DAB, all such atlas frames k in the DAB are removed from the DAB:
   * atlas frame k is marked as "unused for reference".
   * atlas frame k has AtlasFrameOutputFlag equal to 0 or its DAB output time is less than or equal to the CAB removal time of the first decoding unit (denoted as decoding unit m) of the current atlas frame n; i.e., DabOutputTime[ k ] is less than or equal to DuCabRemovalTime[ m ].
   * For each atlas frame that is removed from the DAB, the DAB fullness is decremented by one.
     1. Atlas frame output

The processes specified in this clause happen instantaneously at the CAB removal time of access unit n, AuCabRemovalTime[ n ].

When atlas n has AtlasFrameOutputFlag equal to 1, its DAB output time DabOutputTime[ n ] is derived as follows, where the variable firstAtlasFrameInBufferingPeriodFlag is equal to 1 if access unit n is the first access unit of a buffering period and 0 otherwise:

DabOutputTime[ n ] = AuCabRemovalTime[ n ] + ClockTick \* atlasFrameDafbOutputDelay   
 if( firstAtlasFrameInBufferingPeriodFlag) (E‑11)  
 DabOutputTime[ n ] −= ClockTick \* DabDelayOffset

where atlasFrameDafbOutputDelay is the value of aft\_dab\_output\_delay in the atlas frame timing SEI message associated with access unit n, when present, in the decoding unit information SEI messages associated with access unit n.

The output of the current atlas is specified as follows:

– If AtlasFrameOutputFlag is equal to 1 and DabOutputTime[ n ] is equal to AuCabRemovalTime[ n ], the current atlas frame is output.

– Otherwise, if AtlasFrameOutputFlag is equal to 0, the current atlas frame is not output, but will be stored in the DAB as specified in clause E.3.4

– Otherwise (AtlasFrameOutputFlag is equal to 1 and DabOutputTime[ n ] is greater than AuCabRemovalTime[ n ] ), the current atlas frame is output later and will be stored in the DAB (as specified in clause E.3.4) and is output at time DabOutputTime[ n ] unless indicated not to be output at a time that precedes DabOutputTime[ n ].

When atlas frame n is an atlas frame that is output and is not the last atlas frame of the bitstream that is output, the value of the variable DabOutputInterval[ n ] is derived as follows:

DabOutputInterval[ n ] = DabOutputTime[ nextAtlasFrameInOutputOrder ] −   
 DabOutputTime[ n ] (E‑12)

where nextAtlasFrameInOutputOrder  is the atlas frame that follows atlas frame n in output order and has AtlasFrameOutputFlag equal to 1.

* + 1. Current decoded atlas frame marking and storage

The current decoded atlas frame is stored in the DAB in an empty atlas frame storage buffer, the DAB fullness is incremented by one. When asps\_long\_term\_ref\_atlas\_frames\_flag is equal to 1, this atlas frame is marked as "used for long-term reference". After all the tile groups of the current atlas frame have been decoded, this atlas frame is marked as "used for short-term reference".

NOTE – Unless more memory than required by the level limit is available for storage of decoded atlas frames, decoders should start storing decoded parts of the current atlas frames into the DAB when the first tile group is decoded and continue storing more decoded samples as the decoding process proceeds.

* + 1. Removal of atlas frames from the DAB after decoding of the current atlas frame

Immediately after decoding of the current atlas frame, at the CAB removal time of the last decoding unit of access unit n (containing the current atlas frame), the current decoded atlas frame is removed from the DAB, and the DAB fullness is decremented by one.

* 1. Bitstream conformance

A bitstream of coded data conforming to this document shall fulfil all requirements specified in this clause.

The bitstream shall be constructed according to the syntax, semantics and constraints specified in this document outside of this annex.

The first coded atlas in a bitstream shall be an IRAP atlas, i.e., an IDR atlas frame, or a CRA atlas frame.

The patch bitstream is tested by the HRD for conformance as specified in clause E.1.

For each current atlas frame, let the variables maxAtlasFrameOrderCnt and minAtlasFrameOrderCnt be set equal to the maximum and the minimum, respectively, of the AtlasFrameOrderCntVal values of the following atlas frames:

– The current atlas frame.

– The previous atlas frame in decoding order.

– The short-term reference atlas frames in the reference atlas list (RAL) of the current atlas frame.

– All atlas frames n that have AtlasFrameOutputFlag equal to 1, AuCabRemovalTime[ n ] less than AuCabRemovalTime[ currAtlasFrame ] and DabOutputTime[ n ] greater than or equal to AuCabRemovalTime[ currAtlasFrame ], where currAtlasFrame is the current atlas frame.

All of the following conditions shall be fulfilled for each of the bitstream conformance tests:

1. For each atlas access unit n, with n greater than 0, associated with a buffering period SEI message, let the variable deltaTime90k[ n ] be specified as follows:

deltaTime90k[ n ] = 90000 \* ( AuNominalRemovalTime[ n ] −   
 AuFinalArrivalTime[ n − 1 ] ) (E‑13)

The value of InitCabRemovalDelay[ SchedSelIdx ] is constrained as follows:

– If hrd\_cbr\_flag[ SchedSelIdx ] is equal to 0, the following condition shall be true:

InitCabRemovalDelay[ SchedSelIdx ] <= Ceil( deltaTime90k[ n ] ) (E‑14)

– Otherwise (hrd\_cbr\_flag[ SchedSelIdx ] is equal to 1), the following condition shall be true:

Floor( deltaTime90k[ n ] ) <= InitCabRemovalDelay[ SchedSelIdx ] <=   
 Ceil( deltaTime90k[ n ] ) (E‑15)

NOTE 1 – The exact number of bits in the CAB at the removal time of each atlas frame may depend on which buffering period SEI message is selected to initialize the HRD. Encoders must take this into account to ensure that all specified constraints must be obeyed regardless of which buffering period SEI message is selected to initialize the HRD, as the HRD may be initialized at any one of the buffering period SEI messages.

1. A CAB overflow is specified as the condition in which the total number of bits in the CAB is greater than the CAB size. The CAB shall never overflow.
2. When low\_delay\_hrd\_flag[ HighestTid ] is equal to 0, the CAB shall never underflow. A CAB underflow is specified as follows:

– A CAB underflow is specified as the condition in which the nominal CAB removal time of access unit n AuNominalRemovalTime[ n ] is less than the final CAB arrival time of access unit n AuFinalArrivalTime[ n ] for at least one value of n.

1. The nominal removal times of atlases from the CAB (starting from the second atlas in decoding order) shall satisfy the constraints on AuNominalRemovalTime[ n ] and AuCabRemovalTime[ n ] expressed in Annex A.
2. For each current atlas frame, after invocation of the process for removal of atlass from the DAB as specified in clause E.3.2, the number of decoded atlas frames in the DAB, including all atlas frames n that are marked as "used for reference", or that have AtlasFrameOutputFlag equal to 1 and AuCabRemovalTime[ n ] less than AuCabRemovalTime[ currAtlasFrame ], where currAtlasFrame  is the current atlas frame, shall be less than or equal to asps\_max\_dec\_atlas\_frame\_buffering\_minus1.
3. All reference atlas frames shall be present in the DAB when needed for prediction. Each atlas frame that has AtlasFrameOutputFlag equal to 1 shall be present in the DAB at its DAB output time unless it is removed from the DAB before its output time by one of the processes specified in clause E.3.2.
4. The value of DabOutputInterval[ n ] as given by Equation E-12, which is the difference between the output time of an atlas frame and that of the first atlas frame following it in output order and having AtlasFrameOutputFlag equal to 1, shall satisfy the constraint expressed in Annex A for the profile, tier and level specified in the bitstream using the decoding process specified in clauses 2 through 10.
5. For any two atlass m and n in the same CAS, when DabOutputTime[ m ] is greater than DabOutputTime[ n ], the AtlasFrameOrderCntVal of atlas frame m shall be greater than the AtlasOrderCntVal of atlas n.

NOTE 2 – All atlass of an earlier CAS in decoding order that are output are output before any atlass of a later CAS in decoding order. Within any particular CAS, the atlass that are output are output in increasing AtlasOrderOrderCntVal order.

* 1. Decoder conformance
     1. General

A decoder conforming to this document shall fulfil all requirements specified in this clause.

A decoder claiming conformance to a specific profile, tier and level shall be able to successfully decode all atlas frame bitstreams that conform to the atlas frame bitstream conformance requirements specified in clause E.4, in the manner specified in Annex A, provided that all ASPSs, and AFPSs referred to in the ACL NAL units and appropriate buffering period and atlas frame timing SEI messages are conveyed to the decoder, in a timely manner, either in the bitstream (by non-ACL NAL units), or by external means not specified in this document.

When an atlas frame bitstream contains syntax elements that have values that are specified as reserved and it is specified that decoders shall ignore values of the syntax elements or NAL units containing the syntax elements having the reserved values, and the atlas frame bitstream is otherwise conforming to this document, a conforming decoder shall decode the atlas frame bitstream in the same manner as it would decode a conforming atlas frame bitstream and shall ignore the syntax elements or the NAL units containing the syntax elements having the reserved values as specified.

There are two types of conformance that can be claimed by an atlas frame decoder: output timing conformance and output order conformance.

To check conformance of an atlas frame decoder, test bitstreams conforming to the claimed profile, tier and level, as specified in clause E.4 are delivered by a hypothetical atlas frame stream scheduler (HSS) both to the HRD and to the atlas frame decoder under test (DUT). All decoded atlas frames output by the HRD shall also be output by the DUT, each decoded atlas frame output by the DUT shall be an atlas frame with AtlasFrameOutputFlag equal to 1, and, for each such decoded atlas frame output by the DUT, the values of all atlas units that are output shall be equal to the values of the atlas units produced by the specified atlas frame decoding process.

For output timing atlas frame decoder conformance, the HSS operates as described above, with delivery schedules selected only from the subset of values of SchedSelIdx for which the bit rate and CAB size are restricted as specified in Annex A for the specified profile, tier and level or with "interpolated" delivery schedules as specified below for which the bit rate and CAB size are restricted as specified in Annex A. The same delivery schedule is used for both the HRD and the DUT.

When the HRD parameters and the buffering period SEI messages are present with hrd\_cab\_cnt\_minus1[ HighestTid ] greater than 0, the atlas frame decoder shall be capable of decoding the atlas frame bitstream as delivered from the HSS operating using an "interpolated" delivery schedule specified as having peak bit rate r, CAB size c( r ) and initial CAB removal delay ( f( r ) ÷ r ) as follows:

α = ( r − BitRate[ SchedSelIdx − 1 ] ) ÷ ( BitRate[ SchedSelIdx ] − BitRate[ SchedSelIdx − 1 ] ), (E‑16)

c( r ) = α \* CabSize[ SchedSelIdx ] + ( 1 − α ) \* CabSize[ SchedSelIdx − 1 ] (E‑17)

f( r ) = α \* InitCabRemovalDelay[ SchedSelIdx ] \* BitRate[ SchedSelIdx ] +   
 ( 1 − α ) \* InitCabRemovalDelay[ SchedSelIdx − 1 ] \* BitRate[ SchedSelIdx − 1 ] (E‑18)

for any SchedSelIdx > 0 and r such that BitRate[ SchedSelIdx − 1 ] <= r <= BitRate[ SchedSelIdx ] such that r and c( r ) are within the limits as specified in Annex A for the maximum bit rate and buffer size for the specified profile, tier and level.

NOTE 1 – InitCabRemovalDelay[ SchedSelIdx ] can be different from one buffering period to another and have to be re-calculated.

For output timing atlas frame decoder conformance, an HRD as described above is used and the timing (relative to the delivery time of the first bit) of atlas frame output is the same for both the HRD and the DUT up to a fixed delay.

For output order atlas frame decoder conformance, the following applies:

– The HSS delivers the bitstream AtlasBitstreamToDecode to the DUT "by demand" from the DUT, meaning that the HSS delivers bits (in decoding order) only when the DUT requires more bits to proceed with its processing.

NOTE 2 – This means that for this test, the coded atlas frame buffer of the DUT could be as small as the size of the largest decoding unit.

– A modified HRD as described below is used, and the HSS delivers the bitstream to the HRD by one of the schedules specified in the atlas frame bitstream AtlasBitstreamToDecode such that the bit rate and CAB size are restricted as specified in Annex A. The order of atlas frame output shall be the same for both the HRD and the DUT.

– The HRD CAB size is given by CabSize[ SchedSelIdx ] as specified in clause G.3.3, where SchedSelIdx and the HRD parameters are selected as specified in clause E.1. The DAB size is given by asps\_max\_dec\_atlas\_frame\_buffering\_minus1 + 1. Removal time from the CAB for the HRD is the final bit arrival time and decoding is immediate. The operation of the DAB of this HRD is as described in clauses E.5.2 through E.5.2.3.

* + 1. Operation of the output order DAB
       1. General

The decoded atlas frame buffer contains atlas frame storage buffers. Each of the atlas frame storage buffers contains a decoded atlas frame that is marked as "used for reference" or is held for future output. The process for output and removal of atlas frames from the DAB before decoding of the current atlas frame as specified in clause E.5.2.2 is invoked, the invocation of the process for current decoded atlas frame marking and storage as specified in clause E.5.2.3, further followed by the invocation of the process for removal of atlas frame from the DAB after decoding of the current atlas frame as specified in clause E.5.2.4, and finally followed by the invocation of the process for additional bumping as specified in clause E.5.2.3. The "bumping" process is specified in clause E.5.2.4 and is invoked as specified in clauses  and E.5.2.3.

* + - 1. Output and removal of atlas frames from the DAB

The output and removal of atlas frames from the DAB before the decoding of the current atlas frame (but after parsing the tile group header of the first tile group of the current atlas frame) happens instantaneously when the first decoding unit of the access unit containing the current atlas frame is removed from the CAB and proceeds as follows:

– The decoding process for RAL as specified in clause 8.4.3.2 is invoked.

– If the current atlas is an IRAP atlas frame with NoRaslOutputFlag equal to 1 that is not atlas frame 0, the following ordered steps are applied:

1. The variable NoOutputOfPriorAtlasFramesFlag is derived for the atlas frame decoder under test as follows:

– If the current atlas frame is a CRA atlas frame, NoOutputOfPriorAtlasFramesFlag is set equal to 1.

– Otherwise, NoOutputOfPriorAtlasFramesFlag is set equal to 0.

2. The value of NoOutputOfPriorAtlasFramesFlag derived for the decoder under test is applied for the HRD as follows:

– If NoOutputOfPriorAtlasFramesFlag is equal to 1, all atlas storage buffers in the DAB are emptied without output of the atlas frames they contain and the DAB fullness is set equal to 0.

– Otherwise (NoOutputOfPriorAtlasFramesFlag is equal to 0), all atlas frame storage buffers containing atlas frames that are marked as "not needed for output" and "unused for reference" are emptied (without output) and all non-empty atlas frames storage buffers in the DAB are emptied by repeatedly invoking the "bumping" process specified in clause E.5.2.4 and the DAB fullness is set equal to 0.

– Otherwise (the current atlas frame is not an IRAP atlas frame with NoRaslOutputFlag equal to 1), all atlas frames storage buffers containing atlas frames which are marked as "not needed for output" and "unused for reference" are emptied (without output). For each atlas frame storage buffer that is emptied, the DAB fullness is decremented by one. When one or more of the following conditions are true, the "bumping" process specified in clause E.5.2.4 is invoked repeatedly while further decrementing the DAB fullness by one for each additional atlas frame storage buffer that is emptied, until none of the following conditions is true:

* The number of atlas frames in the DAB is greater than or equal to asps\_max\_dec\_atlas\_frame\_buffering\_minus1 + 1.
  + - 1. Additional bumping

The processes specified in this clause happen instantaneously when the last decoding unit of access unit n containing the current atlas frames is removed from the CAB.

When the current atlas frames has AtlasFrameOutputFlag equal to 1, for each atlas frames in the DAB that is marked as "needed for output" and follows the current atlas frames in output order, the associated variable AtlasFrameLatencyCount is set equal to AtlasFrameLatencyCount + 1.

The following applies:

– If the current decoded atlas has AtlasFrameOutputFlag equal to 1, it is marked as "needed for output" and its associated variable AtlasFrameLatencyCount is set equal to 0.

– Otherwise (the current decoded atlas framehas AtlasFrameOutputFlag equal to 0), it is marked as "not needed for output".

When one or more of the following conditions are true, the "bumping" process specified in clause E.5.2.4 is invoked repeatedly until none of the following conditions are true:

– The number of atlas frames in the DAB that are marked as "needed for output" is greater than asps\_max\_dec\_atlas\_frame\_buffering\_minus1.

* + - 1. "Bumping" process

The "bumping" process consists of the following ordered steps:

1. The atlas frame that is first for output is selected as the one having the smallest value of atlasFrameOrderCntVal of all atlas frames in the DAB marked as "needed for output".

NOTE – For any two atlas frames atlasFrameA and atlasFrameB that belong to the same CAS and are output by the "bumping process", when atlasFrameA is output earlier than atlasFrameB, the value of AtlasFrameOrderCntVal of atlasFrameA is less than the value of AtlasFrameOrderCntVal of atlasFrameB.

1. Supplemental enhancement information

(This annex forms an integral part of this Recommendation | International Standard.)

* 1. General

This annex specifies syntax and semantics for SEI message payloads.

SEI messages assist in processes related to decoding, reconstruction, display, or other purposes. This annex defines two types of SEI messages: essential and non-essential.

Non-essential SEI messages are not required by the decoding process. Conforming decoders are not required to process this information for output order conformance to this document (see Annex A for the specification of conformance).

In clause C.5.2 including its subclauses, specification for presence of non-essential SEI messages is also satisfied when those messages (or some subset of them) are conveyed to decoders (or to the HRD) by other means not specified in this document. When present in the bitstream, non-essential SEI messages shall obey the syntax and semantics specified in clause 7.3.8 and this annex. When the content of a non-essential SEI message is conveyed for the application by some means other than presence within the bitstream, the representation of the content of the SEI message is not required to use the same syntax specified in this annex. For the purpose of counting bits, only the appropriate bits that are actually present in the bitstream are counted.

Essential SEI messages are an integral part of the V-PCC bitstream and should not be removed from the bitstream. The essential SEI messages are categorized into two types:

1. Type-A essential SEI messages: These SEIs contain information required to check bitstream conformance and for output timing decoder conformance. Every V-PCC decoder conforming to point A should not discard any relevant Type-A essential SEI messages and shall consider them for bitstream conformance and for output timing decoder conformance.
2. Type-B essential SEI messages: V-PCC decoders that wish to conform to a particular reconstruction profile as specified in Annex A should not discard any relevant Type-B essential SEI messages and shall consider them for 3D point cloud reconstruction and conformance purposes.

Table XXX lists the essential and non-essential SEI messages. In case of essential SEI messages, the type is also specified.

[Ed. (RJ): A table listing essential and non-essential SEI messages should be inserted over here.]

* 1. SEI payload syntax
     1. General SEI message syntax

|  |  |
| --- | --- |
| sei\_payload( payloadType, payloadSize ) { | **Descriptor** |
| if(( nal\_unit\_type  = =  NAL\_PREFIX\_NSEI )  | |  ( nal\_unit\_type  = =  NAL\_PREFIX\_ESEI )) { |  |
| if( payloadType  = =  0 ) |  |
| buffering\_period( payloadSize ) |  |
| else if( payloadType  = =  1 ) |  |
| atlas\_frame\_timing( payloadSize ) |  |
| else if( payloadType  = =  2 ) |  |
| filler\_payload( payloadSize ) |  |
| else if( payloadType  = =  3 ) |  |
| user\_data\_registered\_itu\_t\_t35( payloadSize ) |  |
| else if( payloadType  = =  4 ) |  |
| user\_data\_unregistered( payloadSize ) |  |
| else if( payloadType  = =  5 ) |  |
| recovery\_point( payloadSize ) |  |
| else if( payloadType  = =  6 ) |  |
| no\_display( payloadSize ) |  |
| else if( payloadType  = =  7 ) |  |
| time\_code( payloadSize ) |  |
| else if( payloadType  = =  8 ) |  |
| regional\_nesting( payloadSize ) |  |
| else if( payloadType  = =  9 ) |  |
| sei\_manifest( payloadSize ) |  |
| else if( payloadType  = =  10 ) |  |
| sei\_prefix\_indication( payloadSize ) |  |
| else if( payloadType  = =  11 ) |  |
| geometry\_transformation\_params( payloadSize ) |  |
| else if( payloadType  = =  12 ) |  |
| attribute\_transformation\_params( payloadSize ) |  |
| else if( payloadType  = =  13 ) |  |
| active\_sub\_bitstreams( payloadSize ) |  |
| else if( payloadType  = =  14 ) |  |
| component\_codec\_mapping( payloadSize ) |  |
| else if( payloadType  = =  15 ) |  |
| volumetric\_tiling\_info( payloadSize ) |  |
| else if( payloadType  = =  16 ) |  |
| presentation\_information( payloadSize ) |  |
| else if( payloadType  = =  17 ) |  |
| geometry\_smoothing( payloadSize ) |  |
| else if( payloadType  = =  18 ) |  |
| attribute\_smoothing( payloadSize ) |  |
| else |  |
| reserved\_sei\_message( payloadSize ) |  |
| } |  |
| else { /\*( nal\_unit\_type  = =  NAL\_SUFFIX\_NSEI )  | |  ( nal\_unit\_type  = =  NAL\_SUFFIX\_ESEI )\*/ |  |
| if( payloadType  = =  2 ) |  |
| filler\_payload( payloadSize ) |  |
| else if( payloadType  = =  3 ) |  |
| user\_data\_registered\_itu\_t\_t35( payloadSize ) |  |
| else if( payloadType  = =  4 ) |  |
| user\_data\_unregistered( payloadSize ) |  |
| else |  |
| reserved\_sei\_message( payloadSize ) |  |
| } |  |
| if( more\_data\_in\_payload( ) ) { |  |
| if( payload\_extension\_present( ) ) |  |
| **sp\_reserved\_payload\_extension\_data** | u(v) |
| byte\_alignment( ) |  |
| } |  |
| } |  |

* + 1. Filler payload SEI message syntax

|  |  |
| --- | --- |
| filler\_payload( payloadSize ) { | Descriptor |
| for( k = 0; k < payloadSize; k++ ) |  |
| **ff\_byte** /\* equal to 0xFF \*/ | f(8) |
| } |  |

* + 1. User data registered by Recommendation ITU-T T.35 SEI message syntax

|  |  |
| --- | --- |
| user\_data\_registered\_itu\_t\_t35( payloadSize ) { | Descriptor |
| **itu\_t\_t35\_country\_code** | b(8) |
| if( itu\_t\_t35\_country\_code  !=  0xFF ) |  |
| i = 1 |  |
| else { |  |
| **itu\_t\_t35\_country\_code\_extension\_byte** | b(8) |
| i = 2 |  |
| } |  |
| do { |  |
| **itu\_t\_t35\_payload\_byte** | b(8) |
| i++ |  |
| } while( i < payloadSize ) |  |
| } |  |

* + 1. User data unregistered SEI message syntax

|  |  |
| --- | --- |
| user\_data\_unregistered( payloadSize ) { | Descriptor |
| **uuid\_iso\_iec\_11578** | u(128) |
| for( i = 16; i < payloadSize; i++ ) |  |
| **user\_data\_payload\_byte** | b(8) |
| } |  |

* + 1. Recovery point SEI message syntax

|  |  |
| --- | --- |
| recovery\_point( payloadSize ) { | Descriptor |
| **recovery\_afoc\_cnt** | se(v) |
| **exact\_match\_flag** | u(1) |
| **broken\_link\_flag** | u(1) |
| } |  |

* + 1. No display SEI message syntax

|  |  |
| --- | --- |
| no\_display( payloadSize ) { | Descriptor |
| } |  |

* + 1. Reserved SEI message syntax

|  |  |
| --- | --- |
| reserved\_sei\_message( payloadSize ) { | Descriptor |
| for( i = 0; i < payloadSize; i++ ) |  |
| **reserved\_sei\_message\_payload\_byte** | b(8) |
| } |  |

* + 1. SEI manifest SEI message syntax

|  |  |
| --- | --- |
| sei\_manifest( payloadSize ) { | **Descriptor** |
| **manifest\_num\_sei\_msg\_types** | u(16) |
| for( i = 0; i < manifest\_num\_sei\_msg\_types; i++ ) { |  |
| **manifest\_sei\_payload\_type**[ i ] | u(16) |
| **manifest\_sei\_description**[ i ] | u(8) |
| } |  |
| } |  |

* + 1. SEI prefix indication SEI message syntax

|  |  |
| --- | --- |
| sei\_prefix\_indication( payloadSize ) { | **Descriptor** |
| **prefix\_sei\_payload\_type** | u(16) |
| **num\_sei\_prefix\_indications\_minus1** | u(8) |
| for( i = 0; i  <=  num\_sei\_prefix\_indications\_minus1; i++ ) { |  |
| **num\_bits\_in\_prefix\_indication\_minus1**[ i ] | u(16) |
| for( j = 0; j  <=  num\_bits\_in\_prefix\_indication\_minus1[ i ]; j++ ) |  |
| **sei\_prefix\_data\_bit**[ i ][ j ] | u(1) |
| while( !byte\_aligned( ) ) |  |
| **byte\_alignment\_bit\_equal\_to\_one** /\* equal to 1 \*/ | f(1) |
| } |  |
| } |  |



* + 1. Attribute transformation parameters SEI message syntax

|  |  |
| --- | --- |
| attribute\_transformation\_params( payloadSize ) { | **Descriptor** |
| **atp\_cancel\_flag** | u(1) |
| if( !atp\_cancel\_flag ) { |  |
| **atp\_num\_attribute\_updates** | ue(v) |
| for( j = 0; j < atp\_num\_attribute\_updates; j++ ) { |  |
| **atp\_attribute\_idx**[ j ] | u(8) |
| **atp\_dimension\_minus1**[atp\_attribute\_idx[ j ]] | u(8) |
| for( i = 0; i < atp\_dimension\_minus1[atp\_attribute\_idx[ j ]]; i++ ) { |  |
| **atp\_scale\_params\_enabled\_flag**[atp\_attribute\_idx[ j ]][ i ] | u(1) |
| **atp\_offset\_params\_enabled\_flag**[atp\_attribute\_idx[ j ]] [ i ] | u(1) |
| if( atp\_scale\_params\_enabled\_flag[atp\_attribute\_idx[ j ]][ i ] ) |  |
| **atp\_attribute\_scale\_minus1**[ atp\_attribute\_idx[ j ] ] [ i ] | u(32) |
| if( atp\_offset\_params\_enabled\_flag[atp\_attribute\_idx[ j ]][ i ] ) |  |
| **atp\_attribute\_offset**[ atp\_attribute\_idx[ j ] ] [ i ] | i(32) |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |

[Ed. (LK) It was discussed to introduce a float descriptor (IEEE-754 single precision format). Could a scale be a float number? could an offset be a negative value?.]

* + 1. Active sub-bitstreams SEI message syntax

|  |  |
| --- | --- |
| active\_sub\_bitstreams( payloadSize ) { | **Descriptor** |
| **asb\_active\_sub\_bitstreams\_cancel\_flag** | u(1) |
| if ( !asb\_active\_sub\_bitstreams\_cancel\_flag ) { |  |
| **asb\_active\_attributes\_changes\_flag** | u(1) |
| **asb\_active\_maps\_changes\_flag** | u(1) |
| **asb\_auxiliary\_sub\_bitstreams\_active\_flag** | u(1) |
| if ( asb\_active\_attributes\_changes\_flag ) { |  |
| **asb\_all\_attributes\_active\_flag** | u(1) |
| if ( !asb\_all\_attributes\_active\_flag ) { |  |
| **asb\_active\_attribute\_count\_minus1** | u(7) |
| for( i = 0; i <= asb\_active\_attribute\_count\_minus1; i++ ) |  |
| **asb\_active\_attribute\_idx**[i] | u(7) |
| } |  |
| } |  |
| if ( asb\_active\_maps\_changes\_flag ) { |  |
| **asb\_all\_maps\_active\_flag** | u(1) |
| if ( !asb\_all\_maps\_active\_flag ) { |  |
| **asb\_active\_map\_count\_minus1** | u(4) |
| for( i = 0; i <= asb\_active\_map\_count\_minus1 ) |  |
| **asb\_active\_map\_idx**[ i ] | u(4) |
| } |  |
| } |  |
| } |  |
| } |  |

* + 1. Component codec mapping SEI message syntax

|  |  |
| --- | --- |
| component\_codec\_mapping( payloadSize ) { | **Descriptor** |
| **ccm\_component\_codec\_cancel\_flag** | u(1) |
| if( !ccm\_component\_codec\_cancel\_flag ) { |  |
| **ccm\_codec\_mappings\_count\_minus1** | u(8) |
| for( i=0; i <= ccm\_codec\_mappings\_count\_minus1; i++ ) { |  |
| **ccm\_codec\_id** | u(8) |
| **ccm\_codec\_4cc**[ ccm\_codec\_id ] | st(v) |
| } |  |
| } |  |
| } |  |

* + 1. Volumetric Annotation SEI message family syntax
       1. Scene Object Information SEI message syntax

|  |  |
| --- | --- |
| scene\_object\_information( payloadSize ) { |  |
| **soi\_cancel\_flag** | u(1) |
| **soi\_num\_object\_updates** | ue(v) |
| if ( soi\_num\_object\_updates > 0 ) { |  |
| **soi\_simple\_objects\_flag** | u(1) |
| if ( soi\_simple\_objects\_flag == 0) { |  |
| **soi\_object\_label\_present\_flag** | u(1) |
| **soi\_priority\_present\_flag** | u(1) |
| **soi\_object\_hidden\_present\_flag** | u(1) |
| **soi\_object\_dependency\_present\_flag** | u(1) |
| **soi\_visibility\_cones\_present\_flag** | u(1) |
| **soi\_3d\_bounding\_box\_present\_flag** | u(1) |
| **soi\_collision\_shape\_present\_flag** | u(1) |
| **soi\_point\_style\_present\_flag** | u(1) |
| **soi\_material\_id\_present\_flag** | u(1) |
| **soi\_extension\_present\_flag** | u(1) |
| } |  |
| else { |  |
| soi\_object\_label\_present\_flag = 0 |  |
| soi\_priority\_present\_flag = 0 |  |
| soi\_object\_hidden\_present\_flag = 0 |  |
| soi\_object\_dependency\_present\_flag = 0 |  |
| soi\_visibility\_cones\_present\_flag = 0 |  |
| soi\_3d\_bounding\_box\_present\_flag = 0 |  |
| soi\_collision\_shape\_present\_flag = 0 |  |
| soi\_point\_style\_present\_flag = 0 |  |
| soi\_material\_id\_present\_flag = 0 |  |
| soi\_extension\_present\_flag = 0 |  |
| } |  |
| if (soi\_3d\_bounding\_box\_present\_flag) { |  |
| **soi\_3d\_bounding\_box\_scale\_log2** | u(5) |
| **soi\_3d\_bounding\_box\_precision\_minus8** | u(5) |
| } |  |
| **soi\_log2\_max\_object\_idx\_updated** | u(5) |
| if( soi\_object\_dependency\_present\_flag ) |  |
| **soi\_log2\_max\_object\_dependency\_idx** | u(5) |
| for( i = 0; i  <=  soi\_num\_object\_updates; i++ ) { |  |
| **soi\_object\_idx**[ i ] | u(v) |
| k = soi\_object\_idx[ i ] |  |
| **soi\_object\_cancel\_flag**[ k ] | u(1) |
| ObjectTracked[ k ] = ! pi\_object\_cancel\_flag[ k ] ] |  |
| if (!soi\_object\_cancel\_flag[ k ]) { |  |
| if( soi\_object\_label\_present\_flag) { |  |
| **soi\_object\_label\_update\_flag**[ k ] | u(1) |
| if( soi\_object\_label\_update\_flag[ k ] ) |  |
| **soi\_object\_label\_idx**[ k ] | ue(v) |
| } |  |
| if( soi\_priority\_present\_flag ) { |  |
| **soi\_priority\_update\_flag**[ k ] | u(1) |
| if( soi\_priority\_update\_flag[ k ] ) |  |
| **soi\_priority\_value**[ k ] | u(4) |
| } |  |
| if( soi\_object\_hidden\_present\_flag ) |  |
| **soi\_object\_hidden\_flag**[ k ] | u(1) |
| if( soi\_object\_dependency\_present\_flag ) { |  |
| **soi\_object\_dependency\_update\_flag**[ k ] | u(1) |
| if (soi\_object\_dependency\_update\_flag[ k ]) { |  |
| **soi\_object\_num\_dependencies**[ k ] | u(4) |
| for( j = 0; j  <  soi\_object\_num\_dependencies[ k ]; j++ ) |  |
| **soi\_object\_dependency\_idx**[ k ] [ j ] | u(v) |
| } |  |
| } |  |
| if( soi\_visibility\_cones\_present\_flag ) { |  |
| **soi\_visibility\_cones \_update\_flag**[ k ] | u(1) |
| if( soi\_visibility\_cones\_update\_flag[ k ]) { |  |
| **soi\_direction\_x**[ k ] | fl(32) |
| **soi\_direction\_y**[ k ] | fl(32) |
| **soi\_direction\_z**[ k ] | fl(32) |
| **soi\_angle**[ k ] | fl(16) |
| } |  |
| } |  |
| if( soi\_3d\_bounding\_box\_present\_flag ) { |  |
| **soi\_3d\_bounding\_box\_update\_flag**[ k ] | u(1) |
| if( soi\_3d\_bounding\_box\_update\_flag[ k ]) { |  |
| **soi\_3d\_bounding\_box\_x**[ k ] | u(v) |
| **soi\_3d\_bounding\_box\_y**[ k ] | u(v) |
| **soi\_3d\_bounding\_box\_z**[ k ] | u(v) |
| **soi\_3d\_bounding\_box\_delta\_x**[ k ] | u(v) |
| **soi\_3d\_bounding\_box\_delta\_y**[ k ] | u(v) |
| **soi\_3d\_bounding\_box\_delta\_z**[ k ] | u(v) |
| } |  |
| } |  |
| if( soi\_collision\_shape\_present\_flag ) { |  |
| **soi\_collision\_shape\_update\_flag**[ k ] | u(1) |
| if (soi\_collision\_shape\_update\_flag[ k] ]) |  |
| **soi\_collision\_shape\_id**[ k ] | u(16) |
| } |  |
| if( soi\_point\_style\_present\_flag ) { |  |
| **soi\_point\_style\_update\_flag**[ k ] | u(1) |
| if (soi\_point\_style\_update\_flag[ k] ]) { |  |
| **soi\_point\_shape\_id**[ k ] | u(8) |
| **soi\_point\_size**[ k ] | u(16) |
| } |  |
| } |  |
| if( soi\_material\_id\_present\_flag ) { |  |
| **soi\_material\_id\_update\_flag**[ k ] | u(1) |
| if (soi\_material\_id\_update\_flag[ k] ]) |  |
| **soi\_material\_id**[ k ] | u(16) |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |

* + - 1. Object Label Information SEI message syntax

|  |  |  |
| --- | --- | --- |
| object\_label\_information ( payload\_size) { | | **Descriptor** |
| **oli\_cancel\_flag** | u(1) | |
| if (!oli\_cancel\_flag) { |  | |
| **oli\_label\_language\_present\_flag** | | u(1) |
| if( oli\_label\_language\_present\_flag ) { | |  |
| while( !byte\_aligned( ) ) | |  |
| **oli\_bit\_equal\_to\_zero** /\* equal to 0 \*/ | | f(1) |
| **oli\_label\_language** | | st(v) |
| } | |  |
| **oli\_num\_label\_updates** | | ue(v) |
| for( i = 0; i < oli\_num\_label\_updates; i++ ) { | |  |
| **oli\_label\_idx**[ i ] | | ue(v) |
| **oli\_label\_cancel\_flag** | | u(1) |
| LabelAssigned[ pi\_label\_idx[ i] ] = !pi\_label\_cancel\_flag | |  |
| if ( !oli\_label\_cancel\_flag) { | |  |
| while( !byte\_aligned( ) ) | |  |
| **oli\_bit\_equal\_to\_zero** /\* equal to 0 \*/ | | f(1) |
| **oli\_label**[ oli\_label\_idx[ i ] ] | | st(v) |
| } | |  |
| } | |  |
| } | |  |
| } | |  |

* + - 1. Patch Information SEI message syntax

|  |  |  |
| --- | --- | --- |
| patch\_information ( payload\_size ) { | | **Descriptor** |
| **pi\_cancel\_flag** | | u(1) |
| **pi\_num\_tile\_group\_updates** | | ue(v) |
| if ( pi\_num\_tile\_group\_updates > 0 ) { | ue(v) | | |
| **pi\_log2\_max\_object\_idx\_tracked** | u(5) | | |
| **pi\_log2\_max\_patch\_idx\_updated** | u(4) | | |
| } |  | | |
| for( i = 0; i  <=  pi\_num\_tile\_group\_updates; i++ ) { | |  |
| **pi\_tile\_group\_address**[ i ] | | ue(v) |
| j = pi\_tile\_group\_address[ i ] | |  |
| **pi\_tile\_group\_cancel\_flag**[ j ] | | u(1) |
| **pi\_num\_patch\_updates**[ j ] | | ue(v) |
| for( k = 0; k  <  pi\_num\_patch\_updates[ j ]; k++ ) { | |  |
| **pi\_patch\_idx**[ j ][ k ] | | u(v) |
| p = pi\_patch\_idx[ j ][ k ] | |  |
| **pi\_patch\_cancel\_flag**[ j ][ p ] | | u(1) |
| if( !pi\_patch\_cancel\_flag[ j ][ p ] ) { | |  |
| **pi\_patch\_number\_of\_objects\_minus1**[ j ][ p ] | | ue(v) |
| m = pi\_patch\_number\_of\_objects\_minus1[ j ][ p ] + 1 | |  |
| for( n = 0; n < m; n++ ) | |  |
| **pi\_patch\_object\_idx**[ j ][ p ][ n ] | | u(v) |
| } | |  |
| } | |  |
| } | |  |
| } | |  |











* + - 1. Volumetric rectangle information SEI message syntax

|  |  |
| --- | --- |
| volumetric\_rectangle\_information ( payload\_size ) { | **Descriptor** |
| **vri\_cancel\_flag** | u(1) |
| **vri\_num\_rectangles\_updates** | ue(v) |
| if ( vri\_num\_rectangles\_updates > 0 ) { | ue(v) |
| **vri\_log2\_max\_object\_idx\_tracked** | u(5) |
| **vri\_log2\_max\_rectangle\_idx\_updated** | u(4) |
| } |  |
| for( k = 0; k  <  vri\_num\_rectangles\_updates]; k++ ) { |  |
| **vri\_rectangle\_idx**[ k ] | u(v) |
| p = vri\_rectangle\_idx[ k ] |  |
| **vri\_rectangle\_cancel\_flag**[ p ] | u(1) |
| if( !vri\_rectangle\_cancel\_flag[ p ] ) { |  |
| **vri\_bounding\_box\_update\_flag**[ p ] | u(1) |
| if( vri\_bounding\_box\_update\_flag[ p ] ) { |  |
| **vri\_bounding\_box\_top**[ p ] | u(v) |
| **vri\_bounding\_box\_left**[ p ] | u(v) |
| **vri\_bounding\_box\_width**[ p ] | u(v) |
| **vri\_bounding\_box\_height**[ p ] | u(v) |
| } |  |
| **vri\_rectangle\_number\_of\_objects\_minus1**[ p ] | ue(v) |
| m = vri\_rectangle\_number\_of\_objects\_minus1[ p ] + 1 |  |
| for( n = 0; n < m; n++ ) |  |
| **vri\_rectangle\_object\_idx**[ p ][ n ] | u(v) |
| } |  |
| } |  |
| } |  |

* + 1. Buffering period SEI message syntax

|  |  |
| --- | --- |
| buffering\_period( payloadSize ) { | **Descriptor** |
| **bp\_atlas\_sequence\_parameter\_set\_id** | ue(v) |
| **bp\_irap\_cab\_params\_present\_flag** | u(1) |
| if( bp\_irap\_cab\_params\_present\_flag ) { |  |
| **bp\_cab\_delay\_offset** | u(v) |
| **bp\_dab\_delay\_offset** | u(v) |
| } |  |
| **bp\_concatenation\_flag** | u(1) |
| **bp\_atlas\_cab\_removal\_delay\_delta\_minus1** | u(v) |
| bp\_max\_sub\_layers\_minus1 = 0 | u(3) |
| for( i = 0; i <= bp\_max\_sub\_layers\_minus1; i++ ) { |  |
| if( NalHrdBpPresentFlag ) { |  |
| for( j = 0; j < hrd\_cab\_cnt\_minus1[ i ] + 1; j++ ) { |  |
| **bp\_nal\_initial\_cab\_removal\_delay**[i][ j ] | u(v) |
| **bp\_nal\_initial\_cab\_removal\_offset**[i][ j ] | u(v) |
| if(bp\_irap\_cab\_params\_present\_flag ) { |  |
| **bp\_nal\_initial\_alt\_cab\_removal\_delay**[ i ] | u(v) |
| **bp\_nal\_initial\_alt\_cab\_removal\_offset**[ i ] | u(v) |
| } |  |
| } |  |
| if( AclHrdBpPresentFlag ) { |  |
| for( j = 0; j < hrd\_cab\_cnt\_minus1[ i ] + 1; j++ ) { |  |
| **bp\_acl\_initial\_cab\_removal\_delay**[ i ][j] | u(v) |
| **bp\_acl\_initial\_cab\_removal\_offset**[ i ][j] | u(v) |
| if(bp\_irap\_cab\_params\_present\_flag ) { |  |
| **bp\_acl\_initial\_alt\_cab\_removal\_delay**[ i ] | u(v) |
| **bp\_acl\_initial\_alt\_cab\_removal\_offset**[ i ] | u(v) |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |

* + 1. Atlas frame timing SEI message syntax

|  |  |
| --- | --- |
| atlas\_frame\_timing( payloadSize ) { | Descriptor |
| if( CabDabDelaysPresentFlag ) { |  |
| **aft\_cab\_removal\_delay\_minus1** | u(v) |
| **aft\_dab\_output\_delay** | u(v) |
| } |  |
| } |  |

* + 1. Presentation information SEI message syntax

|  |  |
| --- | --- |
| presentation\_information( payloadSize ) { | **Descriptor** |
| **pi\_unit\_of\_length** | u(1) |
| **pi\_orientation\_present\_flag** | u(1) |
| **pi\_pivot\_present\_flag** | u(1) |
| **pi\_dimension\_present\_flag** | u(1) |
| if(pi\_orientation\_present\_flag) { |  |
| for( d = 0; d < 3; d++ ) |  |
| **pi\_up**[d] | i(32) |
| **pi\_front**[d] | i(32) |
| } |  |
| if(pi\_pivot\_present\_flag) { |  |
| for( d = 0; d < 3; d++ ) |  |
| **pi\_pivot**[d] | i(64) |
| } |  |
| } |  |
| if(pi\_dimension\_present\_flag) { |  |
| for( d = 0; d < 3; d++ ) |  |
| **pi\_dimension**[d] | i(64) |
| } |  |
| } |  |
| } |  |

[Ed. (LK) up and front could be a float number descriptor if introduced.]

* + 1. Geometry smoothing SEI message syntax

|  |  |
| --- | --- |
| geometry\_smoothing ( payloadSize  ) { | **Descriptor** |
| **gs\_smoothing\_cancel\_flag** | u(1) |
| **gs\_smoothing\_instances\_updated** | u(8) |
| for( i = 0; i < gs\_smoothing\_instances\_updated; i++ ) { |  |
| **gs\_smoothing\_instance\_index**[ i ] | u(8) |
| k = gs\_smoothing\_instance\_index[ i ] |  |
| **gs\_smoothing\_instance\_cancel\_flag**[ k ] | u(1) |
| if( gs\_smoothing\_instance\_cancel\_flag[ k ] != 1 ) { |  |
| **gs\_smoothing\_method\_type**[ k ] | u(8) |
| if (gs\_smoothing\_method\_type[ k ] == 1 ) { | u(1) |
| **gs\_smoothing\_grid\_size\_minus2**[ k ] | u(7) |
| **gs\_smoothing\_threshold**[ k ] | u(8) |
| } |  |
| else if( gs\_smoothing\_method\_type[ k ] == 2 ) { |  |
| **gs\_pbf\_log2\_threshold\_minus1**[ k ] | u(2) |
| **gs\_pbf\_passes\_count\_minus1**[ k ] | u(2) |
| **gs\_pbf\_filter\_size\_minus1**[ k ] | u(3) |
| } |  |
| } |  |
| } |  |
| } |  |

[Ed. (PA): Do we systemitacilly need the term "smoothing" for each syntax element of the geometry smoothing "gs"? smoothing is already present in the "gs" acronym..]

[Ed. (PA) gs\_smoothing\_method\_type may be an ue(v) -- like in HEVC for such syntax element intent-- with bounds provided in semantics rather than an overkill u(8)]

* + 1. Attribute smoothing SEI message syntax

|  |  |
| --- | --- |
| attribute\_smoothing( payloadSize ) { | **Descriptor** |
| **as\_smoothing\_cancel\_flag** | u(1) |
| **as\_num\_attributes\_updated** | ue(v) |
| for( j = 0; j < as\_num\_attributes\_updated; j++ ) { |  |
| **as\_attribute\_idx**[ j ] | u(7) |
| k = as\_attribute\_idx[ j ] |  |
| **as\_attribute\_smoothing\_cancel\_flag**[ k ] |  |
| **as\_smoothing\_instances\_updated**[ k ] | u(8) |
| for( i=0; i < as\_smoothing\_instances\_updated[ k ]; i++ ) { |  |
| **as\_smoothing\_instance\_index**[ k ][ i ] | u(8) |
| m = as\_smoothing\_instance\_index[ k ][ i ] |  |
| **as\_smoothing\_instance\_cancel\_flag**[ k ][ m ] | u(1) |
| if( as\_smoothing\_instance\_cancel\_flag[ k ][ m ] != 1 ) { |  |
| **as\_smoothing\_method\_type**[ k ][ m ] | u(8) |
| if (as\_smoothing\_method\_type[ k ][ m ] == 1) { |  |
| **as\_smoothing\_grid\_size\_minus2**[ k ][ m ] | u(8) |
| **as\_smoothing\_threshold**[ k ][ m ] | u(8) |
| **as\_smoothing\_threshold\_variation**[ k ][ m ] | u(8) |
| **as\_smoothing\_threshold\_difference**[ k ][ m ] | u(8) |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |
| } |  |

[Ed. (RJ): The text in this section seems different from what was adopted, and what is currently in software. More specifically, there is no concept of a filter index for attribute smoothing.]

* 1. SEI payload semantics
     1. General SEI payload semantics

**sp\_reserved\_payload\_extension\_data** shall not be present in bitstreams conforming to this version of this document. However, decoders conforming to this version of this document shall ignore the presence and value of sp\_reserved\_payload\_extension\_data. When present, the length, in bits, of sp\_reserved\_payload\_extension\_data is equal to 8 \* payloadSize − nEarlierBits − nPayloadZeroBits − 1, where nEarlierBits is the number of bits in the sei\_payload( ) syntax structure that precede the sp\_reserved\_payload\_extension\_data syntax element and nPayloadZeroBits is the number of payload\_bit\_equal\_to\_zero syntax elements at the end of the sei\_payload( ) syntax structure.

NOTE 1 – SEI messages with the same value of payloadType are conceptually the same SEI message regardless of whether they are contained in prefix or suffix SEI atlas data units.

The semantics and persistence scope for each SEI message are specified in the semantics specification for each particular SEI message.

NOTE 2 – Persistence information for SEI messages is informatively summarized in Table F-11.

| Table F-11 — Persistence scope of SEI messages (informative) | |
| --- | --- |
| SEI message | Persistence scope |
| Buffering period | The remainder of the bitstream |
| Atlas frame timing | The access unit containing the SEI message |
| Filler payload | The access unit containing the SEI message |
| User data registered by  Rec. ITU-T T.35 | Unspecified |
| User data unregistered | Unspecified |
| Recovery point | Specified by the syntax of the SEI message |
| Decoded point cloud hash | The access unit containing the SEI message |
| No display | The access unit containing the SEI message |
| Time code | The access unit containing the SEI message |
| Regional nesting | Depending on the region-nested SEI messages; each region-nested SEI message has the same persistence scope as if the SEI message was non-region-nested |
| Geometry transformation parameters | Specified by the syntax of the SEI message |
| Attribute transformation parameters | Specified by the syntax of the SEI message |
| Presentation information | Specified by the syntax of the SEI message |
| Geometry smoothing | Specified by the syntax of the SEI message |
| Attribute smoothing | Specified by the syntax of the SEI message |

The values of some SEI message syntax elements, including regional\_nesting\_id, are split into two sets of value ranges, where the first set is specified as "may be used as determined by the application", and the second set is specified as "reserved for future use by ISO/IEC". Applications should be cautious of potential “collisions” of the interpretation for values of these syntax elements belonging to the first set of value ranges. Since different applications might use these IDs having values in the first set of value ranges for different purposes, particular care should be exercised in the design of encoders that generate SEI messages with these IDs having values in the first set of value ranges, and in the design of decoders that interpret SEI messages with these IDs. This document does not define any management for these values. These IDs having values in the first set of value ranges might only be suitable for use in contexts in which "collisions" of usage (i.e., different definitions of the syntax and semantics of an SEI message with one of these IDs having the same value in the first set of value ranges) are unimportant, or not possible, or are managed – e.g., defined or managed in the controlling application or transport specification, or by controlling the environment in which bitstreams are distributed.

In the following subclauses of this annex, the following applies:

– The current SEI message refers to the particular SEI message.

– The current access unit refers to the access unit containing the current SEI message.

In the following subclauses of this annex, when a particular SEI message applies to a set of one or more maps (instead of a set of operation points), i.e., when the payloadType value is not equal to one of 0 (buffering period) and 1 (atlas frame timing), the following applies:

– The semantics apply independently to each unit the particular SEI message applies.

In the following subclauses of this annex, when a particular SEI message applies to a set of one or more operation points (instead of a set of one or more maps), i.e., when the payloadType value is equal to 0 (buffering period) or 1 (atlas frame timing), the following applies:

– The semantics apply independently to each particular operation point of the set of operation points to which the particular SEI message applies.

– The current operation point refers to the particular operation point.

– The terms "access unit" and "CPCS" apply to the bitstream BitstreamToDecode that is the sub-bitstream of the particular operation point.

[Ed. (ALL): All SEIs should be carefully checked and see whether we should be referring to CPCSs or CASs. In general, the importance of CPCSs is higher than that of CASs for a variety of functionalities, and especially for random access]

* + 1. Filler payload SEI message semantics

This SEI message contains a series of payloadSize bytes of value 0xFF, which can be discarded.

**ff\_byte** shall be a byte having the value 0xFF.

* + 1. User data registered by Recommendation ITU-T T.35 SEI message semantics

This SEI message contains user data registered as specified in Recommendation ITU-T T.35, the contents of which are not specified in this document.

**itu\_t\_t35\_country\_code** shall be a byte having a value specified as a country code by Annex A of Recommendation ITU-T T.35.

**itu\_t\_t35\_country\_code\_extension\_byte** shall be a byte having a value specified as a country code by Annex B of Recommendation ITU-T T.35.

**itu\_t\_t35\_payload\_byte** shall be a byte containing data registered as specified in Recommendation ITU-T T.35.

The ITU-T T.35 terminal provider code and terminal provider oriented code shall be contained in the first one or more bytes of the itu\_t\_t35\_payload\_byte, in the format specified by the Administration that issued the terminal provider code. Any remaining itu\_t\_t35\_payload\_byte data shall be data having syntax and semantics as specified by the entity identified by the ITU-T T.35 country code and terminal provider code.

* + 1. User data unregistered SEI message semantics

This SEI message contains unregistered user data identified by a universal unique identifier (UUID), the contents of which are not specified in this document.

**uuid\_iso\_iec\_11578** shall have a value specified as a UUID according to the procedures of Annex A of ISO/IEC 11578:1996.

**user\_data\_payload\_byte** shall be a byte containing data having syntax and semantics as specified by the UUID generator.

* + 1. Recovery point SEI message semantics

The recovery point SEI message assists a decoder in determining when the decoding process will produce acceptable point cloud frames for reconstruction and display after the decoder initiates random access or after the encoder indicates a broken link in the CPCS. When the decoding process is started with the access unit in decoding order associated with the recovery point SEI message, all decoded point cloud frames at or subsequent to the recovery point in output order specified in this SEI message are indicated to be correct or approximately correct in content. Decoded point cloud frames produced by random access at or before the point cloud frame associated with the recovery point SEI message need not be correct in content until the indicated recovery point, and the operation of the decoding process starting at the point cloud frame associated with the recovery point SEI message may contain references to point cloud frames and related video data that might be unavailable for prediction.

In addition, by use of the broken\_link\_flag, the recovery point SEI message can indicate to the decoder the location of some point cloud frames in the bitstream that can result in serious visual artefacts when reconstructed and displayed, because of potentially missing references.

NOTE 1 – The broken\_link\_flag can be used by encoders to indicate the location of a point after which the decoding process for the decoding of some point cloud frames may cause references to information that, though available for use in the decoding process, is not the information that was used for reference when the bitstream was originally encoded (e.g., due to a splicing operation performed during the generation of the bitstream).

When random access is performed to start decoding from the access unit associated with the recovery point SEI message, the decoder operates as if the associated point cloud frame was the first point cloud frame in the bitstream in decoding order, and the variables prevAtlasFrmOrderCntLsb and prevAtlasFrmOrderCntMsb used in derivation of AtlasFrmOrderCntValare both set equal to 0.

NOTE 2 – When HRD information is present in the bitstream, a buffering period SEI message should be associated with the access unit associated with the recovery point SEI message in order to establish initialization of the HRD buffer model after a random access.

[Ed. (AMT): We can maybe skip having the following text that exists in HEVC: "Any SPS or PPS RBSP that is referred to by a atlas frame associated with a recovery point SEI message or by any atlas frame following such a atlas frame in decoding order shall be available to the decoding process prior to its activation, regardless of whether or not the decoding process is started at the beginning of the bitstream or with the access unit, in decoding order, that is associated with the recovery point SEI message." ]

**recovery\_afoc\_cnt** specifies the recovery point of decoded point cloud frames in output order. If there is a point cloud frame pcfA that follows the current point cloud frame (i.e., the point cloud frame associated with the current SEI message) in decoding order in the CPCS and that has AtlasFrmOrderCntValequal to the AtlasFrmOrderCntValof the current point cloud frame plus the value of recovery\_afoc\_cnt, the point cloud frame pcfA is referred to as the recovery point point cloud frame. Otherwise, the first point cloud frame in output order that has AtlasFrmOrderCntValgreater than the AtlasFrmOrderCntValof the current point cloud frame plus the value of recovery\_afoc\_cnt is referred to as the recovery point point cloud frame. The recovery point point cloud frame shall not precede the current point cloud frame in decoding order. All decoded point cloud frames in output order are indicated to be correct or approximately correct in content starting at the output order position of the recovery point point cloud frame. The value of recovery\_afoc\_cnt shall be in the range of −MaxAtlasFrmOrderCntLsb / 2 to MaxAtlasFrmOrderCntLsb / 2 − 1, inclusive.

**exact\_match\_flag** indicates whether decoded point cloud frames at and subsequent to the specified recovery point in output order derived by starting the decoding process at the access unit associated with the recovery point SEI message will be an exact match to the point cloud frames that would be produced by starting the decoding process at the location of a previous IRAP access unit, if any, in the bitstream. The value 0 indicates that the match may not be exact and the value 1 indicates that the match will be exact. When exact\_match\_flag is equal to 1, it is a requirement of bitstream conformance that the decoded point cloud frames at and subsequent to the specified recovery point in output order derived by starting the decoding process at the access unit associated with the recovery point SEI message shall be an exact match to the point cloud frames that would be produced by starting the decoding process at the location of a previous IRAP access unit, if any, in the bitstream.

[Ed. (AMT): We do not have IRAPs here. Also, exact match should be only for the decoded units that are needed to reconstruct the point cloud frame, i.e. patch information and occupancy/geometry/attribute videos. Not at the reconstruction level. Therefore, some of the text here needs to be rephrased.]

When exact\_match\_flag is equal to 0, the quality of the approximation at the recovery point is chosen by the encoding process and is not specified in this document.

**broken\_link\_flag** indicates the presence or absence of a broken link in the atlas data unit stream at the location of the recovery point SEI message and is assigned further semantics as follows:

– If broken\_link\_flag is equal to 1, point cloud frames produced by starting the decoding process at the location of a previous IRAP access unit may contain undesirable visual artefacts to the extent that decoded point cloud frames at and subsequent to the access unit associated with the recovery point SEI message in decoding order should not be reconstructed and displayed until the specified recovery point in output order.

– Otherwise (broken\_link\_flag is equal to 0), no indication is given regarding any potential presence of visual artefacts.

Regardless of the value of the broken\_link\_flag, point cloud frames subsequent to the specified recovery point in output order are specified to be correct or approximately correct in content.

[Ed. (AMT):Need to consider how to handle broken links in this context. Maybe we should consider new proposals since m47962 did not have full specifics for this SEI]

* + 1. No display SEI message semantics

The no display SEI message indicates that the current point cloud frame should not be displayed.

* + 1. Reserved SEI message semantics

The reserved SEI message consists of data reserved for future backward-compatible use by ISO/IEC. It is a requirement of bitstream conformance that bitstreams shall not contain reserved SEI messages until and unless the use of such messages has been specified by ISO/IEC. Decoders shall ignore reserved SEI messages.

* + 1. SEI manifest SEI message semantics

The SEI manifest SEI message conveys information on SEI messages that are indicated as expected (i.e., likely) to be present or not present. Such information may include:

1. The indication that certain types of SEI messages are expected (i.e., likely) to be present (although not guaranteed to be present) in the CPCS.
2. For each type of SEI message that is indicated as expected (i.e., likely) to be present in the CPCS, the degree of expressed necessity of interpretation of the SEI messages of this type.

The degree of necessity of interpretation of an SEI message type may be indicated as "necessary", "unnecessary", or "undetermined".

An SEI message is indicated by the encoder (i.e., the content producer) as being "necessary" when the information conveyed by the SEI message is considered as necessary for interpretation by the decoder or receiving system in order to properly process the content and enable an adequate user experience; it does not mean that the bitstream is required to contain the SEI message in order to be a conforming bitstream. It is at the discretion of the encoder to determine which SEI messages are to be considered as necessary in a particular CPCS. However, it is suggested that some SEI messages should typically be considered as necessary.

1. The indication that certain types of SEI messages are expected (i.e., likely) not to be present (although not guaranteed not to be present) in the CPCS.

The content of an SEI manifest SEI message may, for example, be used by transport-layer or systems-layer processing elements to determine whether the CPCS is suitable for delivery to a receiving and decoding system, based on whether the receiving system can properly process the CPCS to enable an adequate user experience or whether the CPCS satisfies the application needs.

When an SEI manifest SEI message is present in any access unit of a CPCS, an SEI manifest SEI message shall be present in the first access unit of the CPCS. The SEI manifest SEI message persists in decoding order from the current access unit until the end of the CPCS. When there are multiple SEI manifest SEI messages present in a CPCS, they shall have the same content.

**manifest\_num\_sei\_msg\_types** specifies the number of types of SEI messages for which information is provided in the SEI manifest SEI message.

**manifest\_sei\_payload\_type**[ i ] indicates the payloadType value of the i-th type of SEI message for which information is provided in the SEI manifest SEI message. The values of manifest\_sei\_payload\_type[ m ] and manifest\_sei\_payload\_type[ n ] shall not be identical when m is not equal to n.

**manifest\_sei\_description**[ i ] provides information on SEI messages with payloadType equal to manifest\_sei\_payload\_type[ i ] as specified in Table E-12.

Table F-12 — manifest\_sei\_description[ i ] values

|  |  |
| --- | --- |
| **Value** | **Description** |
| 0 | Indicates that there is no SEI message with payloadType equal to manifest\_sei\_payload\_type[ i ] expected to be present in the CPCS. |
| 1 | Indicates that there are SEI messages with payloadType equal to manifest\_sei\_payload\_type[ i ] expected to be present in the CPCS, and these SEI messages are considered as necessary. |
| 2 | Indicates that there are SEI messages with payloadType equal to manifest\_sei\_payload\_type[ i ] expected to be present in the CPCS, and these SEI messages are considered as unnecessary. |
| 3 | Indicates that there are SEI messages with payloadType equal to manifest\_sei\_payload\_type[ i ] expected to be present in the CPCS, and the necessity of these SEI messages is undetermined. |
| 4-255 | Reserved |

The value of manifest\_sei\_description[ i ] shall be in the range of 0 to 3, inclusive, in bitstreams conforming to this version of this document. Other values for manifest\_sei\_description[ i ] are reserved for future use by ISO/IEC. Decoders shall allow the value of manifest\_sei\_description[ i ] greater than or equal to 4 to appear in the syntax and shall ignore all information for payloadType equal to manifest\_sei\_payload\_type[ i ] signalled in the SEI manifest SEI message and shall ignore all SEI prefix indication SEI messages with prefix\_sei\_payload\_type equal to manifest\_sei\_payload\_type[ i ] when manifest\_sei\_description[ i ] is greater than or equal to 4.

* + 1. SEI prefix indication SEI message semantics

The SEI prefix indication SEI message carries one or more SEI prefix indications for SEI messages of a particular value of payloadType. Each SEI prefix indication is a bit string that follows the SEI payload syntax of that value of payloadType and contains a number of complete syntax elements starting from the first syntax element in the SEI payload.

Each SEI prefix indication for an SEI message of a particular value of payloadType indicates that one or more SEI messages of this value of payloadType are expected (i.e., likely) to be present in the CPCS and to start with the provided bit string. A starting bit string would typically contain only a true subset of an SEI payload of the type of SEI message indicated by the payloadType, may contain a complete SEI payload, and shall not contain more than a complete SEI payload. It is not prohibited for SEI messages of the indicated value of payloadType to be present that do not start with any of the indicated bit strings.

These SEI prefix indications should provide sufficient information for indicating what type of processing is needed or what type of content is included. The former (type of processing) indicates decoder-side processing capability, e.g., whether some type of post-filtering process is needed. The latter (type of content) indicates, for example, whether the bitstream contains subtitle captions in a particular language.

The content of an SEI prefix indication SEI message may, for example, be used by transport-layer or systems-layer processing elements to determine whether the CPCS is suitable for delivery to a receiving and decoding system, based on whether the receiving system can properly process the CPCS to enable an adequate user experience or whether the CPCS satisfies the application needs (as determined in some manner by external means outside the scope of this document).

In one example, for user data registered SEI messages that are used to carry captioning information, an SEI prefix indication should include up to at least the language code; and for user data unregistered SEI messages extended for private use, an SEI prefix indication should include up to at least the UUID.

When an SEI prefix indication SEI message is present in any access unit of a CPCS, an SEI prefix indication SEI message shall be present in the first access unit of the CPCS. The SEI prefix indication SEI message persists in decoding order from the current access unit until the end of the CPCS. When there are multiple SEI prefix indication SEI messages present in a CPCS for a particular value of payloadType, they shall have the same content.

**prefix\_sei\_payload\_type** indicates the payloadType value of the SEI messages for which one or more SEI prefix indications are provided in the SEI prefix indication SEI message. When an SEI manifest SEI message is also present for the CPCS, the value of prefix\_sei\_payload\_type shall be equal to one of the manifest\_sei\_payload\_type[ m ] values for which manifest\_sei\_description[ m ] is equal to 1 to 3, inclusive, as indicated by an SEI manifest SEI message that applies to the CPCS.

**num\_sei\_prefix\_indications\_minus1** plus 1 specifies the number of SEI prefix indications.

**num\_bits\_in\_prefix\_indication\_minus1**[ i ] plus 1 specifies the number of bits in the i-th SEI prefix indication.

**sei\_prefix\_data\_bit**[ i ][ j ] specifies the j-th bit of the i-th SEI prefix indication.

The bits sei\_prefix\_data\_bit[ i ][ j ] for j ranging from 0 to num\_bits\_in\_prefix\_indication\_minus1[ i ], inclusive, follow the syntax of the SEI payload with payloadType equal to prefix\_sei\_payload\_type, and contain a number of complete syntax elements starting from the first syntax element in the SEI payload syntax, and may or may not contain all the syntax elements in the SEI payload syntax. The last bit of these bits (i.e., the bit sei\_prefix\_data\_bit[ i ][ num\_bits\_in\_prefix\_indication\_minus1[ i ] ]) shall be the last bit of a syntax element in the SEI payload syntax, unless it is a bit within an itu\_t\_t35\_payload\_byte or user\_data\_payload\_byte.

NOTE – The exception for itu\_t\_t35\_payload\_byte and user\_data\_payload\_byte is provided because these syntax elements may contain externally-specified syntax elements, and the determination of the boundaries of such externally-specified syntax elements is a matter outside the scope of this document.

**byte\_alignment\_bit\_equal\_to\_one** shall be equal to 1.

* + 1. Attribute transformation parameters SEI message semantics

This SEI message provides information that relate to attribute processing and interpretation during reconstruction.

The persistence scope for this SEI message is the remainder of the bitstream (i.e., the signaled attribute transformation parameters SEI message persists until the end of the stream) or when a new attribute transformation parameters SEI message is encountered.

**atp\_cancel\_flag** equal to 1 indicates that the attribute transformation parameters SEI message cancels the persistence of any previous attribute transformation parameters SEI message in output order that applies to the current map. atp\_cancel\_flag equal to 0 indicates that attribute transformation parameters information follows.

**atp\_num\_attribute\_updates** indicates the number of attributes that is to be updated by the current SEI message.

**atp\_attribute\_idx**[ j ] indicates the attribute index of the j-th attribute that is to be updated by the current SEI message.

**atp\_dimension\_minus1**[ j ] plus 1 indicates the number of dimensions that are expected to be associated with the attribute with index j in the current bitstream.

**atp\_offset\_params\_enabled\_flag**[ j ][ i ]equal to 1 indicates that attribute offset parameters are present in the current atlas tile group attribute parameter set. atp\_offset\_params\_enabled\_flag equal to 0 indicates that attribute offset parameters are not present in the current atlas tile group attribute parameter set. When atp\_offset\_params\_enabled\_flag[ j ][ i ] is not present, it shall be inferred to be equal to 0.

**atp\_attribute\_scale\_minus1**[ j ][ i ] plus 1 indicates the value of the scale to be applied to the values of the i‑th dimension of an attribute with index j. The value of atp\_attribute\_scale\_minus1[ j ][ i ] shall be in the range of 0 to 232 – 1, inclusive**.** When atp\_attribute\_scale\_minus1[ j ][ i ] is not present, it shall be inferred to be equal to 1.

**atp\_attribute\_offset**[ j ][ i ] indicates the value of the offset to be added to the values of the i‑th dimension of an attribute with index j. The value of atp\_attribute\_offset[ j ][ i ] shall be in the range of −231 to 231 – 1, inclusive**.** When atp\_attribute\_offset[ j ][ i ] is not present, it shall be inferred to be equal to 0.

[Ed. (AMT): How is clipping performed? Do we need to specify that? Also how about order of operations if scale and offset are specified. I would assume it should be scale \* value + offset, but we have not specified that so someone could interpet this as scale \* (value + offset). Obviously we need to clarify.]

* + 1. Active sub-bitstreams SEI message semantics

This SEI message informs the V-PCC decoder which attributes and/or maps are not available and can therefore be skipped in the decoding process. Similarly, it can also inform the decoder that RAW points are not available. This is done by signaling changes to active attributes/maps and whether RAW points are active or not. An active sub-bitstreams SEI message references a specific V-PCC sequence parameter set. When only a sub-set of the attributes/maps shall be active, the SEI message contains the attribute/map indices for these attributes/maps. A V-PCC decoder shall consider any other attributes/maps in the referenced VPS that are not listed in the SEI message as inactive and skip them in the decoding and/or rendering process.

The persistence scope for this SEI message is the remainder of the bitstream (i.e., the signaled active attributes persist until the end of the stream) or when a new active attributes SEI message is encountered. Previously defined parameters from an earlier SEI message shall persist if not modified and if the value of asb\_active\_sub\_bitstream\_cancel\_flag is not equal to 1. This SEI message should not be ignored by the decoder.

The semantics of the fields of the active sub-bitstreams SEI message are as follows:

**asb\_active\_sub\_bitstream\_cancel\_flag** equal to 1 indicates that all sub-bitstreams are active. asb\_active\_sub\_bitstream\_cancel\_flag equal to 0 indicates that some sub\_bitstreams may not be active.

**asb\_active\_attributes\_changes\_flag** indicates whether there are activation changes for any of the attribute sub-bitstreams of the V-PCC stream. asb\_active\_attributes\_changes\_flag equal to 1 indicates an activation change for at least one attribute and its associated sub-bitstreams. asb\_active\_attributes\_changes\_flag equal to 0 indicates no changes.

**asb\_active\_maps\_changes\_flag** indicates whether there are activation changes for any of the V-PCC map sub-bitstreams. asb\_active\_maps\_changes\_flag equal to 1 indicates an activation change for at least one map and its associated sub-bitstreams. asb\_active\_maps\_changes\_flag equal to 0 indicates no changes.

**asb\_auxiliary\_sub\_bitstreams\_active\_flag** indicates whether the auxiliary sub-bitstreams for the geometry and all attributes are active or not. asb\_auxiliary\_sub\_bitstreams\_active\_flag equal to 1 indicates that such sub-bitstreams are active. asb\_auxiliary\_sub\_bitstreams\_active\_flag equal to 0 indicates that such sub-bitstreams are not active.

**asb\_all\_attributes\_active\_flag** indicates whether all the attributes signaled in the referenced VPS shall be active. asb\_all\_attributes\_active\_flag equal to 1 indicates that all attributes shall be active. asb\_all\_attributes\_active\_flag equal to 0 indicates that only a sub-set of the attributes may be active.

**asb\_active\_attribute\_count\_minus1** plus 1 indicates the number of active attributes signaled in the active sub-bitstreams SEI message.

**asb\_active\_attribute\_idx**[ i ] indicates the attribute index in the V-PCC VPS for the active attribute at index i in the associated SEI message.

**asb\_all\_maps\_active\_flag** indicates whether all the maps signaled in the referenced VPS shall be active. asb\_all\_maps\_active\_flag equal to 1 indicates that all maps shall be active. asb\_all\_maps\_active\_flag equal to 0 indicates that only a sub-set of the maps may be active.

**asb\_active\_map\_count\_minus1** plus 1 indicates the number of active maps signaled in the active sub-bitstreams SEI message.

**asb\_active\_map\_idx**[ i ] indicates the map index in the V-PCC VPS for the active map at index i in the associated SEI message.

* + 1. Component codec mapping SEI message semantics

This SEI message informs the V-PCC decoder of the codec mapping for the codec ids of the component sub-bitstreams signaled in the V-PCC VPS. Each component sub-bitstream codec id is mapped to a specific codec index in a codec lookup table. The codec ids for the component sub-bitstreams in the V-PCC VPS shall be unique. The component codec mapping SEI message shall be used to signal the initial codec mapping to the decoder at the beginning of the V-PCC bitstream as well as signalling updated mappings when the codec of one or more of the V-PCC components sub-bitstreams changes. A V-PCC decoder receiving a component codec mapping SEI message should instantiate new video decoders for the respective component sub-bitstreams signaled in the message.

The persistence scope for this SEI message is the remainder of the bitstream (i.e., the codec changes for the signaled components persist until the end of the stream) or until a new component codec change SEI message is encountered. Only the codec mapping for codec ids specified in the SEI message shall be updated. Previously defined mappings for other codec ids from an earlier SEI message shall persist if not modified and if the value of ccm\_component\_codec\_cancel\_flag is not equal to 1. This SEI message shall not be ignored by the decoder.

When a component codec mapping SEI message is present in any access unit of a CPCS, a component codec mapping SEI message shall be present in the first access unit of the CPCS. The component codec mapping SEI message persists in decoding order from the current access unit until the end of the CPCS. When there are multiple component codec mapping SEI messages present in a CPCS, they shall have the same content.

The semantics of the fields of the component codec mapping SEI message are as follows:

**ccm\_component\_codec\_cancel\_flag** indicates whether the component codec mapping should be reset to the default mapping defined by the CodecGroup profile type for the bitstream.ccm\_component\_codec\_cancel\_flag equal to 1 indicates that the component codec mapping should be reset to the default mapping defined by the CodecGroup profile type for the bitstream. ccm\_component\_codec\_cancel\_flag equal to 0 indicates that the component codec mapping for some components may be updated.

**ccm\_codec\_mappings\_count\_minus1** plus 1 indicates the number of codec mappings that are listed in this SEI message.

**ccm\_codec\_id** is the codec id that is to be mapped to a particular 4CC codec. This codec id may be associated with one or more of the sub-bitstreams in a V-PCC bitstream, as specified within the active V-PCC VPS.

**ccm\_codec\_4cc**[ j ] is the four-character code (4CC) for the codec mapped to the codec id of value j. The codec code shall be an MP4RA registered code.

* + 1. Volumetric annotation SEI messages family syntax
       1. Scene Object information SEI message semantics

This SEI message defines a set of objects that may be present in a volumetric scene, and optionally assigns different properties to these objects. These objects could then potentially be associated with different types of information, including patches and 2D volumetric rectangles that may be defined using the patch information and volumetric rectangle information SEI messages.

The persistence scope for this SEI message is the remainder of the bitstream or until a new scene object information SEI message is encountered. Only the corresponding parameters specified in the SEI message shall be updated. Previously defined parameters from an earlier SEI message shall persist if not modified and if the value of soi\_cancel\_flag is not equal to 1.

[Ed. Note: We need to allow a "temporary" scope version of these SEIs, so as to enable cases where a frame needs to use different statistics from the "global" ones, but not update those statistics (e.g. for a disposable frame). This should be handled with the introduction of an additional flag. A proposal is expected by the editors at the next meeting.]

**soi\_cancel\_flag** equal to 1 indicates that the scene object information SEI message cancels the persistence of any previous scene object information SEI message in output order..

**soi\_num\_object\_updates** indicates the number of objects that are to be updated by the current SEI.

**soi\_simple\_objects\_flag** equal to 1 indicates that no additional information for an updated or newly introduced object will be signalled. soi\_simple\_objects\_flag equal to 0 indicates that additional information for an updated or newly introduced object may be signalled.

**soi\_object\_label\_present\_flag** equal to 1 indicates that object label information is present in the current scene object information SEI message. soi\_object\_label\_present\_flag equal to 0 indicates that object label information is not present.

**soi\_priority\_present\_flag** equal to 1 indicates that priority information is present in the current scene object information SEI message. soi\_priority\_present\_flag equal to 0 indicates that priority information is not present.

**soi\_object\_hidden\_present\_flag** equal to 1 indicates that hidden object information is present in the current scene object information SEI message. soi\_object\_hidden\_present\_flag equal to 0 indicates that hidden object information is not present.

**soi\_object\_dependency\_present\_flag** equal to 1 indicates that object dependency information is present in the current scene object information SEI message. soi\_object\_dependency\_present\_flag equal to 0 indicates that object dependency information is not present.

**soi\_visibility\_cones\_present\_flag** equal to 1 indicates that visibility cones information is present in the current scene object information SEI message. soi\_visibility\_cones\_present\_flag equal to 0 indicates that visibility cones information is not present.

**soi\_3d\_bounding\_box\_present\_flag** equal to 1 indicates that 3D bounding box information is present in the current scene object information SEI message. soi\_3d\_bounding\_box\_present\_flag equal to 0 indicates that 3D bounding box information is not present.

**soi\_collision\_shape\_present\_flag** equal to 1 indicates that collision information is present in the current scene object information SEI message. soi\_collision\_shape\_present\_flag equal to 0 indicates that collision shape information is not present.

**soi\_point\_style\_present\_flag** equal to 1 indicates that point style information is present in the current scene object information SEI message. soi\_point\_style\_present\_flag equal to 0 indicates that point style information is not present.

**soi\_material\_id\_present\_flag** equal to 1 indicates that material ID information is present in the current scene object information SEI message. soi\_material\_id\_present\_flag equal to 0 indicates that material ID information is not present.

**soi\_extension\_present\_flag** equal to 1 indicates that additional extension information shall be present in the current scene object information SEI message. soi\_extension\_present\_flag equal to 0 indicates that additional extension information is not present. It is a requirement of bitstream conformance to this version of the specification that soi\_extension\_present\_flag shall be equal to 0.

**soi\_3d\_bounding\_box\_scale\_log2** indicates the scale to be applied to the 3D bounding box parameters that may be specified for an object.

**soi\_3d\_bounding\_box\_precision\_minus8** plus 8 indicates the precision of the 3D bounding box parameters that may be specified for an object.

**soi\_log2\_max\_object\_idx\_updated** specifies the number of bitsused to signal the value of an object index in the current scene object information SEI message.

**soi\_log2\_max\_object\_dependency\_idx** specifies the number of bitsused to signal the value of a dependency object index in the current scene object information SEI message.

**soi\_object\_idx**[ i ] indicates the object index of the i-th object to be updated. The number of bits used to represent soi\_object\_idx[ i ] is equal to soi\_log2\_max\_object\_idx\_updated. When soi\_object\_idx[ i ] is not present in the bitstream, its value shall be inferred to be equal to 0.

**soi\_object\_cancel\_flag**[ i ] equal to 1 indicates that the object with index equal to i shall be canceled and that the variable ObjectTracked[ i ] shall be set to 0. Furthermore, all of its associated parameters, including the object label, 3D bounding box parameters, priority information, hidden flag, dependency information, visibility cones, collision shapes, point style, and material id, will be reset to their default values. shall also be set equal to 0. soi\_object\_cancel\_flag equal to 0 indicates that the object with index equal to soi\_object\_idx[ i ] shall be updated with information that follows this element and that that the variable ObjectTracked[ i ] shal be set to 1.

**soi\_object\_label\_update\_flag**[ i ] equal to 1 indicates that object label update information is present for an object with object index i. soi\_object\_label\_update\_flag[ i ] equal to 0 indicates that object label update information is not present.

**soi\_object\_label\_idx**[ i ] indicates the label index of an object with index i.

**soi\_priority\_update\_flag**[ i ] equal to 1 indicates that priority update information is present for an object with object index i. soi\_priority\_update\_flag[ i ] equal to 0 indicates that object priority information is not present.

**soi\_priority\_value**[ i ] indicates the priority of an object with index i. The lower the priority value, the higher the priority.

**soi\_object\_hidden\_flag**[ i ] equal to 1 indicates that the object with index i shall be hidden. soi\_object\_hidden\_flag[ i ] equal to 0 indicates that the object with index i shall become present.

**soi\_object\_dependency\_update\_flag**[ i ] equal to 1 indicates that object dependency update information is present for an object with object index i. soi\_object\_dependency\_update\_flag[ i ] equal to 0 indicates that object dependency update information is not present.

**soi\_object\_num\_dependencies**[ i ] indicates the number of dependencies of object with index i.

**soi\_object\_dependency\_idx**[ i ][ j ] indicates the index of the j-th object that has a dependency with the object with object index i.

**soi\_visibility\_cones\_update\_flag**[ i ] equal to 1 indicates that visibility cones update information is present for an object with object index i. soi\_visibility\_cones\_update\_flag[ i ] equal to 0 indicates that visibility cones update information is not present.

**soi\_direction\_x**[ i ] indicates the normalized x-component value of the direction vector for the visibility cone of an object with object index i. The value of soi\_direction\_x[ i ], when not present, is assumed to be equal to 1.0.

**soi\_direction\_y**[ i ] indicates the normalized y-component value of the direction vector for the visibility cone of an object with object index i. The value of soi\_direction\_y[ i ], when not present, is assumed to be equal to 1.0.

**soi\_direction\_z**[ i ] indicates the normalized z-component value of the direction vector for the visibility cone of an object with object index i. The value of soi\_direction\_z[ i ], when not present, is assumed to be equal to 1.0.

**soi\_angle**[ i ] indicates the angle of the visibility cone along the direction vector in degrees. The value of soi\_angle[ i ], when not present, is assumed to be equal to 180°.

**soi\_3d\_bounding\_box\_update\_flag**[ i ] equal to 1 indicates that 3D bounding box information is present for an object with object index i. soi\_3d\_bounding\_box\_update\_flag[ i ] equal to 0 indicates that 3D bounding box information is not present.

**soi\_3d\_bounding\_box\_x**[ i ] indicates the x coordinate value of the origin position of the 3D bounding box of an object with index i. The default value of soi\_3d\_bounding\_box\_x[ i ] is equal to 0.

**soi\_3d\_bounding\_box\_y**[ i ] indicates the y coordinate value of the origin position of the 3D bounding box of an object with index i. The default value of soi\_3d\_bounding\_box\_y[ i ] is equal to 0.

**soi\_3d\_bounding\_box\_z**[ i ] indicates the z coordinate value of the origin position of the 3D bounding box of an object with index i. The default value of soi\_3d\_bounding\_box\_z[ i ] is equal to 0.

**soi\_3d\_bounding\_box\_delta\_x**[ i ]indicates the size of the bounding box on the x axis of an object with index i. The default value of soi\_3d\_bounding\_box\_delta\_x[ i ] is equal to 0.

**soi\_3d\_bounding\_box\_delta\_y**[ i ]indicates the size of the bounding box on the y axis of an object with index i. The default value of soi\_3d\_bounding\_box\_delta\_y[ i ] is equal to 0

**soi\_3d\_bounding\_box\_delta\_z**[ i ]indicates the size of the bounding box on the z axis of an object with index i. The default value of soi\_3d\_bounding\_box\_delta\_z[ i ] is equal to 0

**soi\_collision\_shape\_update\_flag**[ i ] equal to 1 indicates that collision shape update information is present for an object with object index i. soi\_collision\_shape\_update\_flag[ i ] equal to 0 indicates that collision shape update information is not present.

**soi\_collision\_shape\_id**[ i ] indicates the collision shape id of an object with index i.

**soi\_point\_style\_update\_flag**[ i ] equal to 1 indicates that point style update information is present for an object with object index i. soi\_point\_style\_update\_flag[ i ] equal to 0 indicates that point style update information is not present.

**soi\_point\_shape\_id**[ i ] indicates the point shape id of an object with index i. The default value of soi\_point\_shape\_id[ i ] is equal to 0.

**soi\_point\_size**[ i ] indicates the point size of an object with index i. The default value of soi\_point\_size[ i ] is equal to 1.

**soi\_material\_id\_update\_flag**[ i ] equal to 1 indicates that material ID update information is present for an object with object index i. soi\_point\_style\_update\_flag[ i ] equal to 0 indicates that point style update information is not present.

**soi\_material\_id**[ i ] indicates the material ID of an object with index i. The default value of soi\_material\_id[ i ] is equal to 0.

* + - 1. Object label information SEI message semantics

This SEI message defines a set of labels that could be associated with objects in a volumetric scene.

The persistence scope for this SEI message is the remainder of the bitstream or until a new object label information SEI message is encountered. Only the corresponding parameters specified in the SEI message shall be updated. Previously defined parameters from an earlier SEI message shall persist if not modified and if the value of oli\_cancel\_flag is not equal to 1.

**oli\_cancel\_flag** equal to 1 indicates that the object label information SEI message cancels the persistence of any previous object label information SEI message in output order.

**oli\_label\_language\_present\_flag** equal to 1 indicates that label language information is present in the object label information SEI message. oli\_ label\_language\_present\_flag equal to 0 indicates that label language information is not present.

**oli\_bit\_equal\_to\_zero** shall be equal to 0.

**oli\_label\_language** contains a language tag as specified by IETF RFC 5646 followed by a null termination byte equal to 0x00. The length of the oli\_label\_language syntax element shall be less than or equal to 255 bytes, not including the null termination byte.

**oli\_num\_label\_updates** indicates the number of labels that are to be updated by the current SEI.

**oli\_label\_idx**[ i ] indicates the label index of the i-th label to be updated.

**oli\_label\_cancel\_flag** equal to 1 indicates that the label with index equal to oli\_label\_idx[ i ] shall be canceled and set equal to an empty string. oli\_label\_cancel\_flag equal to 0 indicates that the label with index equal to oli\_label\_idx[ i ] shall be updated with information that follows this element.

**oli\_bit\_equal\_to\_zero** shall be equal to 0.

**oli\_label**[ i ] indicates the label of the i-th label. The length of the vti\_label[ i ] syntax element shall be less than or equal to 255 bytes, not including the null termination byte.

* + - 1. Patch information SEI message semantics

This SEI message defines a table that indicates the set of tile groups and corresponding patches that may be present in the current frame. These patches can be assigned to one or more objects that may be present in a volumetric scene and which were previously defined through one or more scene object information SEI messages.

The persistence scope for this SEI message is the remainder of the bitstream or until a new patch information SEI message is encountered. Only the corresponding parameters specified in the SEI message shall be updated. Previously defined parameters from an earlier SEI message shall persist if not modified and if the value of pi\_cancel\_flag is not equal to 1.

**pi\_cancel\_flag** equal to 1 indicates that the patch information SEI message cancels the persistence of any previous patch information SEI message in output order and that all entries in the patch information table shall be removed.

**pi\_num\_tile\_group\_updates** indicates the number of tile groups that are to be updated in the patch information table by the current SEI message.

**pi\_log2\_max\_object\_idx\_tracked** specifies the number of bitsused to signal the value of a tracked object index in the current patch information SEI message.

**pi\_log2\_max\_patch\_idx\_updated** specifies the number of bitsused to signal the value of an updated patch index in the current patch information SEI message.

**pi\_tile\_group\_address**[ i ] specifies the tile group address for the i-th updated tile group in the current SEI message.

**pi\_tile\_group\_cancel\_flag**[ i ] equal to 1 indicates that the tile group with index i shall be reseted and all patches previously assigned to this tile group will be removed. pi\_tile\_group\_cancel\_flag[ i ] equal to 0 indicates that all patches previously assigned to the tile group will index i will be retained.

**pi\_num\_patch\_updates**[ i ] indicates the number of patches that are to be updated by the current SEI within the tile group with index i in the patch information table.

**pi\_patch\_idx**[ i ][ j ] indicates the patch index of the j-th patch in tile group with index i that is to be updated in the patch information table. The number of bits used to represent pi\_patch\_idx[ i ] is equal to pi\_log2\_max\_patch\_idx\_updated. When pi\_patch\_idx[ i ] is not present in the bitstream, its value shall be inferred to be equal to 0.

**pi\_patch\_cancel\_flag**[ i ][ j ] equal to 1 indicates that the patch with index j in tile group with index i shall be removed from the patch information table.

**pi\_patch\_number\_of\_objects\_minus1**[ i ][ j ] indicates the number of objects that are to be associated with the patch with index j in tile group with index i.

**pi\_patch\_object\_idx**[ i ][ j ][ k ] indicates the k-th object index that is associated with the j-th patch in tile group with index i. The number of bits used to represent pi\_patch\_object\_idx[ i ] is equal to pi\_log2\_max\_object\_idx\_tracked. When pi\_patch\_object\_idx[ i ] is not present in the bitstream, its value shall be inferred to be equal to 0.

* + - 1. Volumetric rectangle information SEI message semantics

This SEI message defines a table that indicates a set of volumetric rectangles within an atlas. These volumetric rectangles can be assigned to one or more objects that may be present in a volumetric scene and that were previously defined through one or more scene object information SEI messages.

The persistence scope for this SEI message is the remainder of the bitstream or until a new volumetric rectangle information SEI message is encountered. Only the corresponding parameters specified in the SEI message shall be updated. Previously defined parameters from an earlier SEI message shall persist if not modified and if the value of vri\_cancel\_flag is not equal to 1.

**vri\_cancel\_flag** equal to 1 indicates that the volumetric rectangles information SEI message cancels the persistence of any previous volumetric rectangles information SEI message in output order and that all entries in the volumetric rectangle information table shall be removed.

**vri\_num\_rectangles\_updates** indicates the number of volumetric rectangles that are to be updated by the current SEI.

**vri\_log2\_max\_object\_idx\_tracked** specifies the number of bitsused to signal the value of a tracked object index in the current volumetric rectangle information SEI message.

**vri\_log2\_max\_rectangle\_idx\_updated** specifies the number of bitsused to signal the value of an updated volumetric rectangle index in the current volumetric rectangle information SEI message.

**vri\_rectangle\_idx**[ i ] ] indicates the i-th volumetric rectangle index that is to be updated in the volumetric rectangle information table. The number of bits used to represent vri\_rectangle\_idx[ i ] is equal to vri\_log2\_max\_rectangle\_idx\_updated. When vri\_rectangle\_idx[ i ] is not present in the bitstream, its value shall be inferred to be equal to 0.

**vri\_rectangle\_cancel\_flag**[ i ] equal to 1 indicates that the volumetric rectangle with index i shall be removed from the volumetric rectangle information table.

**vri\_bounding\_box\_update\_flag**[ i ] equal to 1 indicates that 2D bounding box information for the volumetric rectangle with index i should be updated. vti\_bounding\_box\_update\_flag[ i ] equal to 0 indicates that 2D bounding box information for the volumetric rectangle with index i should not be updated.

**vri\_bounding\_box\_top**[ i ] indicates the vertical coordinate value of the top-left position of the bounding box of the i-th volumetric rectangle within the current atlas frame. The default value of vri\_bounding\_box\_top[ i ] is equal to 0.

**vri\_bounding\_box\_left**[ i ] indicates the horizonal coordinate value of the top-left position of the bounding box of the i-th volumetric rectangle within the current atlas frame. The default value of vri\_bounding\_box\_left[ i ] is equal to 0.

**vri\_bounding\_box\_width**[ i ] indicates the width of the bounding box of the i-th volumetric rectangle. The default value of vri\_bounding\_box\_width[ i ] is equal to 0.

**vri\_bounding\_box\_height**[ i ] indicates the height of the bounding box of the i-th volumetric rectangle. The default value of vri\_bounding\_box\_height[ i ] is equal to 0.

**vri\_rectangle\_number\_of\_objects\_minus1**[ i ] indicates the number of objects that are to be associated with the i-th volumetric rectangle.

**vri\_rectangle\_object\_idx**[ i ][ j ] indicates the j-th object index that is associated with the i-th volumetric rectangle. The number of bits used to represent vri\_rectangle\_object\_idx[ i ] is equal to vri\_log2\_max\_object\_idx\_tracked. When vri\_rectangle\_object\_idx[ i ] is not present in the bitstream, its value shall be inferred to be equal to 0.

* + 1. Buffering period SEI message semantics

A buffering period SEI message provides initial CAB removal delay and initial CAB removal delay offset information for initialization of the HRD at the position of the associated access unit in decoding order.

The following applies for the buffering period SEI message syntax and semantics:

– The syntax elements hrd\_initial\_cab\_removal\_delay\_length\_minus1, hrd\_au\_cab\_removal\_delay\_length\_minus1, hrd\_dab\_output\_delay\_length\_minus1, and the variables NalHrdBpPresentFlag and AclHrdBpPresentFlag are found in or derived from syntax elements found in the hrd\_parameters( ) syntax structure that is applicable to at least one of the operation points to which the buffering period SEI message applies.

– The variables CabSize[ i ], BitRate[ i ] and CabCnt are derived from syntax elements found in the sub\_layer\_hrd\_parameters( ) syntax structure that is applicable to at least one of the operation points to which the buffering period SEI message applies.

– Any two operation points that the buffering period SEI message applies to having different OpTid values tIdA and tIdB indicate that the values of hrd\_cab\_cnt\_minus1[ tIdA ] and hrd\_cab\_cnt\_minus1[ tIdB ] coded in the hrd\_parameters( ) syntax structure(s) applicable to the two operation points are identical.

– Any two operation points that the buffering period SEI message applies to having different OpLayerIdList values layerIdListA and layerIdListB indicate that the values of nal\_hrd\_parameters\_present\_flag and acl\_hrd\_parameters\_present\_flag, respectively, for the two hrd\_parameters( ) syntax structures applicable to the two operation points are identical.

– The bitstream (or a part thereof) refers to the bitstream subset (or a part thereof) associated with any of the operation points to which the buffering period SEI message applies.

The presence of buffering period SEI messages for an operation point is specified as follows:

– If NalHrdBpPresentFlag is equal to 1 or AclHrdBpPresentFlag is equal to 1, the following applies for each access unit in the CAS:

* If the access unit is an IRAP access unit, a buffering period SEI message applicable to the operation point shall be associated with the access unit.
* Otherwise, if both of the following conditions apply, a buffering period SEI message applicable to the operation point may or may not be present for the access unit:
  + The atlas frame has TemporalId equal to 0.
  + The atlas frame is not a RASL, RADL or sub-layer non-reference atlas frame.
* Otherwise, the access unit shall not be associated with a buffering period SEI message applicable to the operation point.

– Otherwise (NalHrdBpPresentFlag and AclHrdBpPresentFlag are both equal to 0), no access unit in the CAS shall be associated with a buffering period SEI message applicable to the operation point.

NOTE 1 – For some applications, frequent presence of buffering period SEI messages may be desirable (e.g., for random access at an IRAP atlas frame or a non-IRAP atlas frame or for bitstream splicing).

**bp\_seq\_parameter\_set\_id** indicates and shall be equal to the asps\_seq\_parameter\_set\_id for the ASPS that is active for the coded atlas frame associated with the buffering period SEI message. The value of bp\_seq\_parameter\_set\_id shall be equal to the value of afps\_seq\_parameter\_set\_id in the AFPS referenced by the atgh\_atlas\_frame\_parameter\_set\_id of the atlas tile group headers of the coded atlas associated with the buffering period SEI message. The value of bp\_seq\_parameter\_set\_id shall be in the range of 0 to 15, inclusive.

**bp\_irap\_cab\_params\_present\_flag** equal to 1 specifies the presence of the bp\_acl\_initial\_alt\_cab\_removal\_delay[ i ] and bp\_acl\_initial\_alt\_cab\_removal\_offset[ i ] syntax elements. When not present, the value of bp\_irap\_cab\_params\_present\_flag is inferred to be equal to 0. When the associated atlas is neither a CRA atlas nor a BLA atlas, the value of bp\_irap\_cab\_params\_present\_flag shall be equal to 0.

**bp\_cab\_delay\_offset** specifies an offset to be used in the derivation of the nominal CAB removal times of access units following, in decoding order, the CRA or BLA access unit associated with the buffering period SEI message when the RASL access units associated with the CRA or BLA access unit are not present. The syntax element has a length in bits given by hrd\_au\_cab\_removal\_delay\_length\_minus1 + 1. When not present, the value of bp\_cab\_delay\_offset is inferred to be equal to 0.

**bp\_dab\_delay\_offset** specifies an offset to be used in the derivation of the DAB output times of the CRA or BLA access unit associated with the buffering period SEI message when the RASL access units associated with the CRA or BLA access unit are not present. The syntax element has a length in bits given by hrd\_dab\_output\_delay\_length\_minus1 + 1. When not present, the value of bp\_dab\_delay\_offset is inferred to be equal to 0.

When the current atlas is not the first atlas in the bitstream in decoding order, let prevNonDiscardableAtlas be the preceding atlas in decoding order with TemporalId equal to 0 that is not a RASL, RADL or sub-layer non-reference atlas frame.

**bp\_concatenation\_flag** indicates, when the current atlas is not the first atlas in the bitstream in decoding order, whether the nominal CAB removal time of the current atlas is determined relative to the nominal CAB removal time of the preceding atlas with a buffering period SEI message or relative to the nominal CAB removal time of the atlas prevNonDiscardablePic.

**bp\_atlas\_cab\_removal\_delay\_delta\_minus1** plus 1, when the current atlas is not the first atlas in the bitstream in decoding order, specifies a CAB removal delay increment value relative to the nominal CAB removal time of the atlas prevNonDiscardablePic. This syntax element has a length in bits given by bp\_atlas\_cab\_removal\_delay\_length\_minus1 + 1.

When the current atlas contains a buffering period SEI message and bp\_concatenation\_flag is equal to 0 and the current atlas is not the first atlas in the bitstream in decoding order, it is a requirement of bitstream conformance that the following constraint applies:

* If the atlas prevNonDiscardableAtlas is not associated with a buffering period SEI message, the aft\_cab\_removal\_delay\_minus1 of the current atlas shall be equal to the aft\_cab\_removal\_delay\_minus1 of prevNonDiscardableAtlas plus bp\_atlas\_cab\_removal\_delay\_delta\_minus1 + 1.
* Otherwise, aft\_cab\_removal\_delay\_minus1 shall be equal to bp\_atlas\_cab\_removal\_delay\_delta\_minus1.

NOTE 2 – When the current atlas contains a buffering period SEI message and bp\_concatenation\_flag is equal to 1, the aft\_cab\_removal\_delay\_minus1 for the current atlas is not used. The above-specified constraint can, under some circumstances, make it possible to splice bitstreams (that use suitably-designed referencing structures) by simply changing the value of bp\_concatenation\_flag from 0 to 1 in the buffering period SEI message for an IRAP atlas at the splicing point. When bp\_concatenation\_flag is equal to 0, the above-specified constraint enables the decoder to check whether the constraint is satisfied as a way to detect the loss of the atlas prevNonDiscardableAtlas.

**bp\_max\_sub\_layers\_minus1** plus 1 specifies the maximum number of temporal sub-layers. The value of bp\_max\_sub\_layers\_minus1 shall be equal to 0.

**bp\_nal\_initial\_cab\_removal\_delay**[ i ] and **bp\_nal\_initial\_alt\_cab\_removal\_delay**[ i ] specify the default and the alternative initial CAB removal delays, respectively, for the i-th CAB when the NAL HRD parameters are in use. The syntax elements have a length in bits given by hrd\_initial\_cab\_removal\_delay\_length\_minus1 + 1, and are in units of a 90 kHz clock. The values of the syntax elements shall not be equal to 0 and shall be less than or equal to 90000 \* ( CabSize[ i ] ÷ BitRate[ i ] ), the time-equivalent of the CAB size in 90 kHz clock units.

**bp\_nal\_initial\_cab\_removal\_offset**[ i ] and **bp\_nal\_initial\_alt\_cab\_removal\_offset**[ i ] specify the default and the alternative initial CAB removal offsets, respectively, for the i-th CAB when the NAL HRD parameters are in use. The syntax elements have a length in bits given by hrd\_initial\_cab\_removal\_delay\_length\_minus1 + 1 and are in units of a 90 kHz clock.

Over the entire CAS, the sum of bp\_nal\_initial\_cab\_removal\_delay[ i ] and bp\_nal\_initial\_cab\_removal\_offset[ i ] shall be constant for each value of i, and the sum of bp\_nal\_initial\_alt\_cab\_removal\_delay[ i ] and bp\_nal\_initial\_alt\_cab\_removal\_offset[ i ] shall be constant for each value of i.

**bp\_acl\_initial\_cab\_removal\_delay**[ i ] and **bp\_acl\_initial\_alt\_cab\_removal\_delay**[ i ] specify the default and the alternative initial CAB removal delays, respectively, for the i-th CAB when the VCL HRD parameters are in use. The syntax elements have a length in bits given by hrd\_initial\_cab\_removal\_delay\_length\_minus1 + 1, and are in units of a 90 kHz clock. The values of the syntax elements shall not be equal to 0 and shall be less than or equal to 90000 \* ( CabSize[ i ] ÷ BitRate[ i ] ), the time-equivalent of the CAB size in 90 kHz clock units.

**bp\_acl\_initial\_cab\_removal\_offset**[ i ] and **bp\_acl\_initial\_alt\_cab\_removal\_offset**[ i ] specify the default and the alternative initial CAB removal offsets, respectively, for the i-th CAB when the VCL HRD parameters are in use. The syntax elements have a length in bits given by hrd\_initial\_cab\_removal\_delay\_length\_minus1 + 1 and are in units of a 90 kHz clock.

Over the entire CAS, the sum of bp\_acl\_initial\_cab\_removal\_delay[ i ] and bp\_acl\_initial\_cab\_removal\_offset[ i ] shall be constant for each value of i, and the sum of bp\_acl\_initial\_alt\_cab\_removal\_delay[ i ] and bp\_acl\_initial\_alt\_cab\_removal\_offset[ i ] shall be constant for each value of i.

NOTE 3 – Encoders are recommended not to include irap\_cab\_params\_present\_flag equal to 1 in buffering period SEI messages associated with a CRA or BLA atlas for which at least one of its associated RASL atlass follows one or more of its associated RADL atlass in decoding order.

* + 1. Atlas frame timing SEI message semantics

The atlas frame timing SEI message provides CAB removal delay and DAB output delay information for the access unit associated with the SEI message.

The following applies for the atlas timing SEI message syntax and semantics:

– The syntax elements and variable hrd\_au\_cab\_removal\_delay\_length\_minus1, hrd\_dab\_output\_delay\_length\_minus1 and CabDabDelaysPresentFlag are found in or derived from syntax elements found in the hrd\_parameters( ) syntax structure that is applicable to at least one of the operation points to which the atlas timing SEI message applies.

– The bitstream (or a part thereof) refers to the bitstream subset (or a part thereof) associated with any of the operation points to which the atlas timing SEI message applies.

NOTE 1 – The syntax of the atlas timing SEI message is dependent on the content of the hrd\_parameters( ) syntax structures applicable to the operation points to which the atlas timing SEI message applies. These hrd\_parameters( ) syntax structures are in the ASPS that are active for the coded atlas associated with the atlas timing SEI message. When the atlas timing SEI message is associated with an IRAP access unit with NoRaslOutputFlag equal to 1, unless it is preceded by a buffering period SEI message within the same access unit, the activation of the SPS (and, for IRAP atlass with NoRaslOutputFlag equal to 1 that are not the first atlas in the bitstream in decoding order, the determination that the coded atlas is an IRAP NoRaslOutputFlag equal to 1) does not occur until the decoding of the first coded tile group NAL unit of the coded atlas. Since the coded tile group NAL unit of the coded atlas follows the atlas timing SEI message in NAL unit order, there may be cases in which it is necessary for a decoder to store the RBSP containing the atlas timing SEI message until determining the active VPS and the active SPS for the coded atlas, and then perform the parsing of the atlas timing SEI message.

**aft\_cab\_removal\_delay\_minus1** plus 1 specifies the number clock ticks between the nominal CAB removal time of the access unit associated with the atlas timing SEI message and the preceding access unit in decoding order that contained a buffering period SEI message. This value is also used to calculate an earliest possible time of arrival of access unit atlas data into the CAB for the HSS. The syntax element is a fixed length code whose length in bits is given by hrd\_au\_cab\_removal\_delay\_length\_minus1 + 1.

NOTE 2 – The value of hrd\_au\_cab\_removal\_delay\_length\_minus1 that determines the length (in bits) of the syntax element aft\_cab\_removal\_delay\_minus1 is the value of hrd\_au\_cab\_removal\_delay\_length\_minus1 coded in the ASPS that is active for the coded atlas frame associated with the atlas frame timing SEI message, although aft\_cab\_removal\_delay\_minus1 specifies a number of clock ticks relative to the removal time of the preceding atlas access unit containing a buffering period SEI message, which may be an access unit of a different CAS.

The variable AuCabRemovalDelayMsb of the current atlas frame is derived as follows:

* If the current atlas frame is associated with a buffering period SEI message that is applicable to at least one of the operation points to which the atlas frame timing SEI message applies, AuCabRemovalDelayMsb is set equal to 0.
* Otherwise, the following applies:
* Let maxCabRemovalDelay be equal to 2hrd\_au\_cab\_removal\_delay\_length\_minus1 + 1.
* Let prevAuCabRemovalDelayMinus1 and prevAuCabRemovalDelayMsb be set equal to au\_cab\_removal\_delay\_minus1 and AuCabRemovalDelayMsb, respectively, of the previous atlas frame in decoding order that has TemporalId equal to 0, that is not a RASL, RADL or sub-layer non-reference atlas frame, and that is within the same buffering period as the current atlas frame.
* AuCabRemovalDelayMsb is derived as follows:

if( au\_cab\_removal\_delay\_minus1 <= prevAuCabRemovalDelayMinus1 )  
 AuCabRemovalDelayMsb = prevAuCabRemovalDelayMsb + maxCabRemovalDelay (E‑1)  
else  
 AuCabRemovalDelayMsb = prevAuCabRemovalDelayMsb

The variable AuCabRemovalDelayVal is derived as follows:

AuCabRemovalDelayVal = AuCabRemovalDelayMsb + au\_cab\_removal\_delay\_minus1 + 1 (E‑2)

The value of AuCabRemovalDelayVal shall be in the range of 1 to 232, inclusive. Within one buffering period, the AuCabRemovalDelayVal values for any two access units shall not be the same.

**aft\_dab\_output\_delay** is used to compute the DAB output time of the atlas. It specifies how many clock ticks to wait after removal of the access unit from the CAB before the decoded atlas is output from the DAB.

NOTE 3 – A atlas is not removed from the DAB at its output time when it is still marked as "used for short-term reference" or "used for long-term reference".

The length of the syntax element aft\_dab\_output\_delay is given in bits by hrd\_dab\_output\_delay\_length\_minus1 + 1. When asps\_max\_dec\_atlas\_buffering\_minus1[ minTid ] is equal to 0, where minTid is the minimum of the OpTid values of all operation points the atlas timing SEI message applies to, aft\_dab\_output\_delay shall be equal to 0.

The output time derived from the aft\_dab\_output\_delay of any atlas that is output from an output timing conforming decoder shall precede the output time derived from the aft\_dab\_output\_delay of all atlases in any subsequent CAS in decoding order.

The atlas output order established by the values of this syntax element shall be the same order as established by the values of AtlasOrderCntVal.

For atlases that are not output by the "bumping" process because they precede, in decoding order, an IRAP atlas with NoRaslOutputFlag equal to 1, the output times derived from aft\_dab\_output\_delay shall be increasing with increasing value of AtlasOrderCntVal relative to all atlases within the same CAS.

* + 1. Presentation information SEI message semantics

This SEI message provides information on how the reconstructed point cloud should be presented to a user. All the signalling is in respect to a point cloud on which the geometry transformation signalled by geometry transformation parameters SEI message has been already performed. If a geometry transformation parameters SEI message is not present, then geometry transformation should be performed using the default values.

[Ed. AMT: The last sentence sounds as if there is a default transformation (i.e. identity of course). Do we want to say that? Maybe one could modify the previous sentence and just say "This SEI message is associated with the point cloud after any transformation process signalled by any present geometry transformation parameters SEI message is applied". Note btw, what happens if there are multiple such SEIs (both for presentation and geometry transformation). Would it have been better if the two were "linked" together, i.e. attached onto a single SEI? I guess the next sentence handles this one, i.e. persistence is until a new presentation information SEI is encountered, but maybe we should consider if we wish to have separate parameters depending on different transformation SEIs?]

The persistence scope for this SEI message is the remainder of the bitstream (i.e., the signaled presentation information SEI message persists until the end of the stream) or when a new presentation information SEI message is encountered

**pi\_unit\_of\_length\_flag** equal to 1 indicates that the unit of length in a point cloud is a meter. unit\_of\_length\_flag equal to 0 indicates that the unit of length is arbitrary.

[Ed. AMT: Should this be in the VUI? Also, would it be better if this was more generic? Currently this seems too restrictive and it is not clear how to best use this. Also, given that the point cloud data use integer values, would the use of a "meter" really be too coarse?"

**pi\_orientation\_present\_flag** equal to 1 indicates that point cloud orientation information is present in the presentation information SEI message. pi\_orientation\_present\_flag equal to 0 indicates that point cloud orientation information is not present.

**pi\_pivot\_present\_flag** equal to 1 indicates that point cloud pivot point information is present in the presentation information SEI message. pi\_pivot\_present\_flag equal to 0 indicates that point cloud pivot point information is not present.

**pi\_dimension\_present\_flag** equal to 1 indicates that information about the dimension of a space occupied by point cloud sequence is present in the presentation information SEI message. pi\_dimension\_present\_flag equal to 0 indicates that information about the dimension of a space occupied by the point cloud sequence is not present. [Ed. AMT: sequence? dimension of a space?]

**pi\_up**[ d ] indicates a value along the d axis of a unit vector that describes the up direction of the reconstructed point cloud sequence in units of 2-16. When pi\_up[ d ] is not present, it shall be inferred to represent a unit vector equal to (0.0, 0.0, 1.0). [Ed. AMT: Given the representation, the default vector should be equal to (0, 0, 65536). The same applies to the other similar elements] [Ed. AMT: sequence? This could be changed per frame so it seems not appropriate to talk about sequences]

**pi\_front**[ d ] indicates a value along the d axis of a unit vector that describes the front direction of the reconstructed point cloud sequence in units of 2-16. When pi\_front[ d ] is not present, it shall be inferred to represent a unit vector equal to (0.0, 1.0, 0.0).

**pi\_pivot**[ d ] indicates a position of a pivot point along the d axis in units of 2-32. When pi\_pivot[ d ] is not present, it shall be inferred to be equal to represent a pivoit point equal to (0.0, 0.0, 0.0).

**pi\_dimension**[ d ] indicates a dimension along the d axis of a space occupied by the point cloud sequence in units of 2-32. [Ed. AMT: What happens if not present?]

* + 1. Geometry smoothing SEI message semantics

This SEI message specifies the recommended geometry smoothing methods and their associated parameters that could be used to process the geometry of the current point cloud frame after it is reconstructed, so as to obtain improved reconstructed geometry quality.

Up to 256 different geometry smoothing instances could be specified for use with each point cloud frame. These instances are indicated using an array GeometrySmoothingMethod. The geometry smoothing instance that a decoder may select to operate in, is outside the scope of this document.

At the start of each sequence, let GeometrySmoothingMethod[ i ] = 0, where i corresponds to the geometry smoothing instance index and is in the range of 0 to 255, inclusive. When GeometrySmoothingMethod[ i ] is equal to 0 it means that no geometry smoothing filter is indicated for the geometry smoothing instance with index i.

**gs\_smoothing\_cancel\_flag** equal to 1 indicates that the geometry smoothing SEI message cancels the persistence of any previous geometry smoothing SEI message in output order and resets all entries in the array GeometrySmoothingMethod to 0.

**gs\_smoothing\_instances\_updated** specifies the number of geometry smoothing instances that will be updated in the current geometry smoothing SEI message.

**gs\_smoothing\_instance\_index**[ i ] indicates the i-th geometry smoothing instance index in the array GeometrySmoothingMethod that is to be updated by the current SEI message.

**gs\_smoothing\_instance\_cancel\_flag**[ k ]equal to 1 indicates that the value of GeometrySmoothingMethod[ k ] for the geometry smoothing instance with index k should be set to 0.

**gs\_smoothing\_method\_type**[ k ]indicates the geometry smoothing method that can be used for processing the current point cloud frame as specified in Table E.13 for geometry smoothing instance with index k.

Table F.13 – Definition of gs\_smoothing\_method\_type[ k ]

|  |  |
| --- | --- |
| **Value** | **Interpretation** |
| 0 | No geometry smoothing |
| 1 | Grid based geometry smoothing |
| 2 | Patch border filtering |
| 3-255 | Reserved |

Values of gs\_smoothing\_method\_type[ k ] in the range of 3 to 255, inclusive, are reserved for future use by ISO/IEC. It is a requirement of bitstream conformance that bitstreams conforming to this version of this document shall not contain such values of gs\_smoothing\_method\_type[ k ]. Decoders shall ignore geometry smoothing SEI messages that contain reserved values of gs\_smoothing\_method\_type[ k ].

**gs\_smoothing\_grid\_size\_minus2**[ k ] plus 2 specifies the value of the variable GeometrySmoothingGridSize[ k ] used when processing the current point cloud frame for geometry smoothing instance with index k. The value of gs\_smoothing\_grid\_size\_minus2[ k ] shall be in the range of 0 to 127, inclusive. The value of GeometrySmoothingGridSize[ k ] is computed as follows:

GeometrySmoothingGridSize[ k ] = gs\_smoothing\_grid\_size\_minus2[ k ] + 2

**gs\_smoothing\_threshold**[ k ] indicates the smoothing threshold used for processing the current point cloud frame for geometry smoothing instance with index k. The value of gs\_smoothing\_threshold[ k ] shall be in the range of 0 to 255, inclusive.

**gs\_pbf\_log2\_threshold\_minus1**[ k ] specifies the value of the variable PbfThreshold[ k ] used for processing the current point cloud frame for geometry smoothing instance with index k when the patch border filtering process is used. The value of PbfThreshold[ k ] is computed as follows:

PbfThreshold[ k ] = 2gs\_pbf\_log2\_threshold\_minus1[ k ] + 1

The value of gs\_pbf\_log2\_threshold\_minus1[ k ] shall be in the range of 0 to 3, inclusive.

**gs\_pbf\_passes\_count\_minus1**[ k ]plus 1 specifies the value of the variable PbfPasses[ k ] used for processing the current point cloud frame for geometry smoothing instance with index k when the patch border filtering process is used. The value of gs\_pbf\_passes\_count\_minus1[ k ] shall be in the range of 0 to 3, inclusive.

**gs\_pbf\_filter\_size\_minus1**[ k ]plus 1 specifies the values of the variables PbfSizeX[ k ] and PbfSizeY[ k ] used for processing the current point cloud frame for geometry smoothing instance with index k when the patch border filtering process is used. The values of PbfSizeX[ k ] and PbfSizeY( k ] are computed as follows:

PbfSizeX[ k ] = gs\_pbf\_filter\_size\_minus1[ k ] + 1

PbfSizeY[ k ] = PbfSizeX[ k ] >> 1

The value of gs\_pbf\_filter\_size\_minus1[ k ]shall be in the range of 0 to 7, inclusive.

* + 1. Attribute smoothing SEI message semantics

This SEI message specifies the recommended attribute smoothing methods and any associated parameters that could be used to process the attributes of the current point cloud frame after it is reconstructed, so as to obtain improved reconstructed attribute quality.

For each attribute, up to 256 different attribute smoothing instances could be specified for use with each point cloud frame. These instances are indicated using a two dimensional array AttributeSmoothingMethod. The attribute smoothing instance that a decoder may select to operate in, is outside the scope of this document.

[Ed. (RJ): The text in this section seems different from what was adopted, and what is currently in software. More specifically, there is no concept of a filter index for attribute smoothing.]

At the start of each sequence, let AttributeSmoothingMethod[ j ][ i ] = 0 for all j and i, where j corresponds to the attribute index and is in the range of 0 to 127, inclusive, and i corresponds to the attribute smoothing instance index and is in the range of 0 to 255, inclusive. When AttributeSmoothingMethod[ j ][ i ] is equal to 0 it means that no attribute smoothing filter is indicated for an attribute index j and its corresponding attribute smoothing instance with index i.

**as\_smoothing\_cancel\_flag** equal to 1 indicates that the attribute smoothing SEI message cancels the persistence of any previous attribute smoothing SEI message in output order and resets all entries in the array AttributeSmoothingMethod to 0.

**as\_num\_attributes\_updated** specifies the number of attributes for which attribute smoothing instances will be updated in the current attribute smoothing SEI message.

**as\_attribute\_idx**[ j ]specifies the j-th attribute index that is to be updated.

**as\_attribute\_smoothing\_cancel\_flag**[ k ]equal to 1 indicates that all attribute smoothing instances associated with the attribute with index k should be set to 0, and that AttributeSmoothingMethod[ k ][ i ], for i being in the range of 0 to 255, should be set to 0.

**as\_smoothing\_instances\_updated**[ k ]specifies the number of attribute smoothing instances that will be updated for the attribute with index k.

**as\_smoothing\_instance\_index**[ k ][ i ]specifies the i-th index of the attribute smoothing instance to be updated for attribute with index k.

**as\_smoothing\_instance\_cancel\_flag**[ k ][ m ]equal to 1 indicates that the attribute smoothing instance with index m associated with the attribute with index k should be set to 0, and that AttributeSmoothingMethod[ k ][ m ] should be set to 0.

**as\_smoothing\_method\_type**[ k ][ m ]indicates the attribute smoothing method that is associated with the m-th attribute smoothing instance of attribute with index k as specified in Table E.14

Table E.14 – Definition of as\_smoothing\_method\_type[ k ][ m ]

|  |  |
| --- | --- |
| **Value** | **Interpretation** |
| 0 | No attribute smoothing |
| 1 | Grid based attribute smoothing |
| 2-255 | Reserved |

Values of as\_smoothing\_method\_type[ k ][ m ] in the range of 2 to 255, inclusive, are reserved for future use by ISO/IEC. It is a requirement of bitstream conformance that bitstreams conforming to this version of this document shall not contain such values of as\_smoothing\_method\_type[ k ][ m ]. Decoders shall ignore attribute smoothing SEI messages that contain reserved values of as\_smoothing\_method\_type[ k ][ m ].

**as\_smoothing\_grid\_size\_minus2**[ k ][ m ] plus 2 specifies the value of the variable AttributeSmoothingGridSize[ k ][ m ] used for attribute with index k that is associated with attribute smoothing instance with index m. The value of as\_smoothing\_grid\_size\_minus2[ k ][ m ] shall be in the range of 0 to 127, inclusive. When not present, the value of as\_smoothing\_grid\_size\_minus2[ k ][ m ] is inferred to be equal to 0. The value of AttributeSmoothingGridSize[ k ][ m ] is computed as follows:

AttributeSmoothingGridSize[ k ][ m ] = as\_smoothing\_grid\_size\_minus2[ k ][ i ] + 2

**as\_smoothing\_threshold**[ k ][ m ] indicates the attribute smoothing threshold for attribute smoothing instance with index m of attribute with index k. The value of as\_smoothing\_threshold[ k ][ m ] shall be in the range of 0 to 255, inclusive.

**as\_smoothing\_threshold\_variation**[ k ][ m ] indicates the threshold of attribute variation for attribute smoothing instance with index m of attribute with index k. The value of as\_smoothing\_threshold\_variation[ k ][ m ] shall be in the range of 0 to 255, inclusive.

**as\_smoothing\_threshold\_difference**[ k ][ m ]indicates the threshold of attribute difference for attribute smoothing instance with index m of attribute with index k. The value of as\_smoothing\_threshold\_difference[ k ][ m ] shall be in the range of 0 to 255, inclusive.

[Ed. All: For some of these variables, does a value of 0 make sense? If not, i.e. if it results in no change of an attribute, then that seems a wasteful indication and maybe we should consider either disallowing it implicitly or explicitly]

1. Volumetric usability information

(This annex forms an integral part of this Recommendation | International Standard.)

* 1. General

This annex specifies syntax and semantics of the VUI parameters of the ASPSs.

VUI parameters are not required for constructing the atlas by the decoding process. Conforming decoders are not required to process this information for output order conformance to this document (see Annex D for the specification of output order conformance). Some VUI parameters are required to check bitstream conformance and for output timing decoder conformance.

In Annex F, specification for presence of VUI parameters is also satisfied when those parameters (or some subset of them) are conveyed to decoders (or to the HRD) by other means not specified in this document. When present in the bitstream, VUI parameters shall follow the syntax and semantics specified in this annex. When the content of VUI parameters is conveyed for the application by some means other than presence within the bitstream, the representation of the content of the VUI parameters is not required to use the same syntax specified in this annex. For the purpose of counting bits, only the appropriate bits that are actually present in the bitstream are counted.

* 1. VUI syntax
     1. VUI parameters syntax

|  |  |
| --- | --- |
| vui\_parameters( ) { | **Descriptor** |
| **vui\_timing\_info\_present\_flag** | u(1) |
| if( vui\_timing\_info\_present\_flag ) { |  |
| **vui\_num\_units\_in\_tick** | u(32) |
| **vui\_time\_scale** | u(32) |
| **vui\_poc\_proportional\_to\_timing\_flag** | u(1) |
| if( vui\_poc\_proportional\_to\_timing\_flag ) |  |
| **vui\_num\_ticks\_poc\_diff\_one\_minus1** | ue(v) |
| **vui\_hrd\_parameters\_present\_flag** | u(1) |
| if( vui\_hrd\_parameters\_present\_flag ) |  |
| hrd\_parameters( ) |  |
| } |  |
| } |  |

* + 1. HRD parameters syntax

|  |  |
| --- | --- |
| hrd\_parameters( ) { | Descriptor |
| **hrd\_nal\_parameters\_present\_flag** | u(1) |
| **hrd\_acl\_parameters\_present\_flag** | u(1) |
| if( hrd\_nal\_parameters\_present\_flag | | hrd\_acl\_parameters\_present\_flag ){ |  |
| **hrd\_bit\_rate\_scale** | u(4) |
| **hrd\_cab\_size\_scale** | u(4) |
| **hrd\_initial\_cab\_removal\_delay\_length\_minus1** | u(5) |
| **hrd\_au\_cab\_removal\_delay\_length\_minus1** | u(5) |
| **hrd\_dab\_output\_delay\_length\_minus1** | u(5) |
| } |  |
| maxNumSubLayersMinus1 = 0 |  |
| for( i = 0; i <= maxNumSubLayersMinus1; i++ ) { |  |
| **hrd\_fixed\_atlas\_rate\_general\_flag**[ i ] | u(1) |
| if( !hrd\_fixed\_atlas\_rate\_general\_flag[ i ] ) |  |
| **hrd\_fixed\_atlas\_rate\_within\_cas\_flag**[ i ] | u(1) |
| if( hrd\_fixed\_atlas\_rate\_within\_cas\_flag[ i ] ) |  |
| **hrd\_elemental\_duration\_in\_tc\_minus1**[ i ] | ue(v) |
| else |  |
| **hrd\_low\_delay\_flag**[ i ] | u(1) |
| if( !hrd\_low\_delay\_flag[ i ] ) |  |
| **hrd\_cab\_cnt\_minus1**[ i ] | ue(v) |
| if( hrd\_nal\_parameters\_present\_flag ) |  |
| hrd\_sub\_layer\_parameters( 0, i ) |  |
| if( hrd\_acl\_parameters\_present\_flag ) |  |
| hrd\_sub\_layer\_parameters( 1,  i ) |  |
| } |  |
| } |  |

* + 1. Sub-layer HRD parameters syntax

|  |  |
| --- | --- |
| hrd\_sub\_layer\_parameters( type, subLayerId ) { | **Descriptor** |
| for( i = 0; i <= CabCnt; i++ ) { |  |
| **hrd\_bit\_rate\_value\_minus1**[ type ][ i ] | ue(v) |
| **hrd\_cab\_size\_value\_minus1**[ type ][ i ] | ue(v) |
| **hrd\_cbr\_flag**[ type ][ i ] | u(1) |
| } |  |
| } |  |

* 1. VUI semantics
     1. VUI parameters semantics

**vui\_timing\_info\_present\_flag** equal to 1 specifies that vui\_num\_units\_in\_tick, vui\_time\_scale, vui\_poc\_proportional\_to\_timing\_flag, and vui\_hrd\_parameters\_present\_flag are present in the vui\_parameters( ) syntax structure. vui\_timing\_info\_present\_flag equal to 0 specifies that vui\_num\_units\_in\_tick, vui\_time\_scale, vui\_poc\_proportional\_to\_timing\_flag, and vui\_hrd\_parameters\_present\_flag are not present in the vui\_parameters( ) syntax structure.

**vui\_num\_units\_in\_tick** is the number of time units of a clock operating at the frequency vui\_time\_scale Hz that corresponds to one increment (called a clock tick) of a clock tick counter. vui\_num\_units\_in\_tick shall be greater than 0. A clock tick, in units of seconds, is equal to the quotient of vui\_num\_units\_in\_tick divided by vui\_time\_scale. For example, when the frame rate of an atlas signal is 25 Hz, vui\_time\_scale may be equal to 27 000 000 and vui\_num\_units\_in\_tick may be equal to 1 080 000 and consequently a clock tick may be equal to 0.04 seconds.

**vui\_time\_scale** is the number of time units that pass in one second. For example, a time coordinate system that measures time using a 27 MHz clock has a vui\_time\_scale of 27 000 000. The value of vui\_time\_scale shall be greater than 0.

**vui\_poc\_proportional\_to\_timing\_flag** equal to 1 indicates that the atlas frame order count value for each atlas in the CAS that is not the first atlas in the CAS, in decoding order, is proportional to the output time of the atlas relative to the output time of the first atlas in the CAS. vui\_poc\_proportional\_to\_timing\_flag equal to 0 indicates that the atlas frame order count value for each atlas in the CAS that is not the first atlas in the CAS, in decoding order, may or may not be proportional to the output time of the atlas relative to the output time of the first atlas in the CAS.

**vui\_num\_ticks\_poc\_diff\_one\_minus1** plus 1 specifies the number of clock ticks corresponding to a difference of atlas frame order count values equal to 1. The value of vui\_num\_ticks\_poc\_diff\_one\_minus1 shall be in the range of 0 to 232 − 2, inclusive.

**vui\_hrd\_parameters\_present\_flag** equal to 1 specifies that the syntax structure hrd\_parameters( ) is present in the vui\_parameters( ) syntax structure. vui\_hrd\_parameters\_present\_flag equal to 0 specifies that the syntax structure hrd\_parameters( ) is not present in the vui\_parameters( ) syntax structure.

* + 1. HRD parameters semantics

The hrd\_parameters( ) syntax structure provides HRD parameters used in the HRD operations for a layer set. It is a requirement for bitstream conformance to this version of the Specification that only one layer shall be supported, i.e. nal\_layer\_id and nal\_temporal\_id shall be equal to 0. When the hrd\_parameters( ) syntax structure is included in an ASPS, the layer set to which the hrd\_parameters( ) syntax structure applies is the layer set for which the associated layer identifier list contains all nal\_layer\_id values present in the CAS.

For interpretation of the following semantics, the bitstream (or a part thereof) refers to the bitstream subset (or a part thereof) associated with the layer set to which the hrd\_parameters( ) syntax structure applies.

**hrd\_nal\_parameters\_present\_flag** equal to 1 specifies that NAL HRD parameters (pertaining to Type II bitstream conformance) are present in the hrd\_parameters( ) syntax structure. hrd\_nal\_parameters\_present\_flag equal to 0 specifies that NAL HRD parameters are not present in the hrd\_parameters( ) syntax structure.

NOTE 1 – When hrd\_nal\_parameters\_present\_flag is equal to 0, the conformance of the bitstream cannot be verified without provision of the NAL HRD parameters and all buffering period and atlas timing SEI messages, by some means not specified in this document.

The variable NalHrdBpPresentFlag is derived as follows:

– If one or more of the following conditions are true, then the value of NalHrdBpPresentFlag is set equal to 1:

– hrd\_nal\_parameters\_present\_flag is present in the bitstream and is equal to 1.

– The need for presence of buffering periods for NAL HRD operation to be present in the bitstream in buffering period SEI messages is determined by the application, by some means not specified in this document.

– Otherwise, the value of NalHrdBpPresentFlag is set equal to 0.

**hrd\_acl\_parameters\_present\_flag** equal to 1 specifies that ACL HRD parameters (pertaining to all bitstream conformance types) are present in the hrd\_parameters( ) syntax structure. hrd\_acl\_parameters\_present\_flag equal to 0 specifies that ACL HRD parameters are not present in the hrd\_parameters( ) syntax structure.

NOTE 2 – When hrd\_acl\_parameters\_present\_flag is equal to 0, the conformance of the bitstream cannot be verified without provision of the ACL HRD parameters and all buffering period and atlas timing SEI messages, by some means not specified in this document.

The variable AclHrdBpPresentFlag is derived as follows:

– If one or more of the following conditions are true, then the value of AclHrdBpPresentFlag is set equal to 1:

– hrd\_acl\_parameters\_present\_flag is present in the bitstream and is equal to 1.

– The need for presence of buffering periods for ACL HRD operation to be present in the bitstream in buffering period SEI messages is determined by the application, by some means not specified in this document.

– Otherwise, the value of AclHrdBpPresentFlag is set equal to 0.

The variable CabDabDelaysPresentFlag is derived as follows:

– If one or more of the following conditions are true, then the value of CabDabDelaysPresentFlag is set equal to 1:

– hrd\_nal\_parameters\_present\_flag is present in the bitstream and is equal to 1.

– hrd\_acl\_parameters\_present\_flag is present in the bitstream and is equal to 1.

– The need for presence of CAB and DAB output delays to be present in the bitstream in atlas timing SEI messages is determined by the application, by some means not specified in this document.

– Otherwise, the value of CabDabDelaysPresentFlag is set equal to 0.

**hrd\_initial\_cab\_removal\_delay\_length\_minus1** plus 1 specifies the length, in bits, of the bp\_nal\_initial\_cab\_removal\_delay[ i ], bp\_nal\_initial\_cab\_removal\_offset[ i ], bp\_acl\_initial\_cab\_removal\_delay[ i ], and bp\_acl\_initial\_cab\_removal\_offset[ i ] syntax elements of the buffering period SEI message. When the hrd\_initial\_cab\_removal\_delay\_length\_minus1 syntax element is not present, it is inferred to be equal to 23.

**hrd\_au\_cab\_removal\_delay\_length\_minus1** plus 1 specifies the length, in bits, of the cab\_delay\_offset syntax element in the buffering period SEI message and the aft\_cab\_removal\_delay\_minus1 syntax element in the atlas timing SEI message. When the hrd\_au\_cab\_removal\_delay\_length\_minus1 syntax element is not present, it is inferred to be equal to 23.

**hrd\_dab\_output\_delay\_length\_minus1** plus 1 specifies the length, in bits, of the bp\_dab\_delay\_offset syntax element in the buffering period SEI message and the aft\_dab\_output\_delay syntax element in the atlas timing SEI message. When the hrd\_dab\_output\_delay\_length\_minus1 syntax element is not present, it is inferred to be equal to 23.

**hrd\_fixed\_atlas\_rate\_general\_flag**[ i ] equal to 1 indicates that, when HighestTid is equal to i, the temporal distance between the HRD output times of consecutive atlases in output order is constrained as specified below. hrd\_fixed\_atlas\_rate\_general\_flag[ i ] equal to 0 indicates that this constraint may not apply.

When hrd\_fixed\_atlas\_rate\_general\_flag[ i ] is not present, it is inferred to be equal to 0.

**hrd\_fixed\_atlas\_rate\_within\_cas\_flag**[ i ] equal to 1 indicates that, when HighestTid is equal to i, the temporal distance between the HRD output times of consecutive atlases in output order is constrained as specified below. hrd\_fixed\_atlas\_rate\_within\_cas\_flag[ i ] equal to 0 indicates that this constraint may not apply.

When hrd\_fixed\_atlas\_rate\_general\_flag[ i ] is equal to 1, the value of hrd\_fixed\_atlas\_rate\_within\_cas\_flag[ i ] is inferred to be equal to 1.

**hrd\_elemental\_duration\_in\_tc\_minus1**[ i ] plus 1 (when present) specifies, when HighestTid is equal to i, the temporal distance, in clock ticks, between the elemental units that specify the HRD output times of consecutive atlases in output order as specified below. The value of hrd\_elemental\_duration\_in\_tc\_minus1[ i ] shall be in the range of 0 to 2047, inclusive.

For each atlas n that is output and is not the last atlas in the bitstream (in output order) that is output, the value of the variable DabOutputElementalInterval[ n ] is specified by:

DabOutputElementalInterval[ n ] = DabOutputInterval[ n ] ÷ DeltaToDivisor (F‑1)

where DabOutputInterval[ n ] is specified in D-12 and DeltaToDivisor is equal to 1.

When HighestTid is equal to i and hrd\_fixed\_atlas\_rate\_general\_flag[ i ] is equal to 1 for a CAS containing atlas n, the value computed for DabOutputElementalInterval[ n ] shall be equal to ClockTick \* ( hrd\_elemental\_duration\_in\_tc\_minus1[ i ] + 1 ), wherein ClockTick is as specified in Equation D-1 (using the value of ClockTick for the CAS containing atlas n) when one of the following conditions is true for the following atlas in output order nextAtlasFrameInOutputOrder that is specified for use in Equation D-12:

– the atlas frame nextAtlasFrameInOutputOrder is in the same CAS as the atlas frame n.

– the atlas frame nextAtlasFrameInOutputOrder is in a different CAS and hrd\_fixed\_atlas\_rate\_general\_flag[ i ] is equal to 1 in the CAS containing the atlas frame nextAtlasFrameInOutputOrder, the value of ClockTick is the same for both CASs, and the value of hrd\_elemental\_duration\_in\_tc\_minus1[ i ] is the same for both CASs.

When HighestTid is equal to i and hrd\_fixed\_atlas\_rate\_within\_cas\_flag[ i ] is equal to 1 for a CAS containing atlas frame n, the value computed for DabOutputElementalInterval[ n ] shall be equal to ClockTick \* ( hrd\_elemental\_duration\_in\_tc\_minus1[ i ] + 1 ), wherein ClockTick is as specified in Equation (using the value of ClockTick for the CAS containing atlas frame n) when the following atlas frame in output order nextAtlasFrameInOutputOrder that is specified for use in Equation 10 is in the same CAS as atlas frame n.

**hrd\_low\_delay\_flag**[ i ] specifies the HRD operational mode, when HighestTid is equal to i, as specified in Annex D. When not present, the value of hrd\_low\_delay\_flag[ i ] is inferred to be equal to 0.

NOTE 3 – When hrd\_low\_delay\_flag[ i ] is equal to 1, "big atlases" that violate the nominal CAB removal times due to the number of bits used by an access unit are permitted. It is expected, but not required, that such "big atlases" occur only occasionally.

**hrd\_cab\_cnt\_minus1**[ i ] plus 1 specifies the number of alternative CAB specifications in the bitstream of the CAS when HighestTid is equal to i. The value of hrd\_cab\_cnt\_minus1[ i ] shall be in the range of 0 to 31, inclusive. When not present, the value of hrd\_cab\_cnt\_minus1[ i ] is inferred to be equal to 0.

* + 1. HRD sub-layer parameters semantics

The variable CabCnt is set equal to hrd\_cab\_cnt\_minus1[ subLayerId ].

**hrd\_bit\_rate\_value\_minus1**[ j ][ i ] (together with hrd\_bit\_rate\_scale) specifies the maximum input bit rate for the i-th CAB when the CAB operates at the access unit level for a particular HRD type j. hrd\_bit\_rate\_value\_minus1[ j ][ i ] shall be in the range of 0 to 232 − 2, inclusive. For any i > 0, hrd\_bit\_rate\_value\_minus1[ j ][ i ] shall be greater than hrd\_bit\_rate\_value\_minus1[ j ][ i − 1 ].

The bit rate in bits per second is given by:

BitRate[ j ][ i ] = ( hrd\_bit\_rate\_value\_minus1[ j ][ i ] + 1 ) \* 2( 6 + hrd\_bit\_rate\_scale ) (F‑2)

When the hrd\_bit\_rate\_value\_minus1[ j ][ i ] syntax element is not present, the value of BitRate[ j ][ i ] is inferred to be equal to CalBrAclFactor \* MaxBR for ACL HRD parameters, i.e. when j is equal to 0, and equal to CalBrNalFactor \* MaxBR for NAL HRD parameters, i.e. when j is equal to 1, where MaxBR, CalBrAclFactor and CalBrNalFactor are specified in Annex A.

**hrd\_cab\_size\_value\_minus1**[ j ][ i ] is used together with hrd\_cab\_size\_scale to specify the i-th CAB size when the CAB operates at the access unit level for a particular HRD type j. hrd\_cab\_size\_value\_minus1[ j ][ i ] shall be in the range of 0 to 232 − 2, inclusive. For any i greater than 0, hrd\_cab\_size\_value\_minus1[ j ][ i ] shall be less than or equal to hrd\_cab\_size\_value\_minus1[ j ][ i − 1 ].

The CAB size in bits is given by:

CabSize[ j ][ i ] = ( hrd\_cab\_size\_value\_minus1[ j ][ i ] + 1 ) \* 2( 4 + hrd\_cab\_size\_scale ) (F‑3)

When the hrd\_cab\_size\_value\_minus1[ j ][ i ] syntax element is not present, the value of CabSize[ j ][ i ] is inferred to be equal to CalBrAclFactor \* MaxCAB for ACL HRD parameters, i.e. when j is equal to 1, and equal to CalBrNalFactor \* MaxCAB for NAL HRD parameters, i.e. when j is equal to 0, where MaxCAB, CalBrAclFactor and CalBrNalFactor are specified in Annex A.

**hrd\_cbr\_flag**[ j ][ i ] equal to 0 specifies that to decode this CAS by the HRD type j using the i-th CAB specification, the hypothetical stream scheduler (HSS) operates in an intermittent bit rate mode. hrd\_cbr\_flag[ j ][ i ] equal to 1 specifies that the HSS operates in a constant bit rate (CBR) mode. When not present, the value of hrd\_cbr\_flag[ j ][ i ] is inferred to be equal to 0.

Bibliography

1. Recommendation ITU-T H.222.0 (in force), *Information technology – Generic coding of moving pictures and associated audio information: Systems.*

ISO/IEC 13818-1(in force), *Information technology – Generic coding of moving pictures and associated audio information – Part 1: Systems.*

1. Recommendation ITU-T H.320 (in force), *Narrow-band visual telephone systems and terminal equipment*.
2. ISO/IEC 14496-10: *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*.
3. ISO/IEC 23008-2: *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding*.
4. Registration authority for code-points in "MP4 Family" files: https://mp4ra.org/#