

Title: Coding of Parameter Sets

Status: Input Document to JVT

Purpose: Proposal

**Author(s) or
Contact(s):** Miska M. Hannuksela
Nokia Mobile Software
P.O. Box 68
33721 Tampere
Finland

Tel: +358 40 521 2845
Email: miska.hannuksela@nokia.com

Ye-Kui Wang
TICSP
Tampere Univ. of Technology
P.O. Box 553
FIN-33101 Tampere
Finland

+358 3 3115 3972
wyk@ieee.org

Source: Nokia Corporation and Tampere University of Technology

1. SUMMARY

In the JVT NAL design, infrequently changing information is organized into the parameter set structure introduced in VCEG-N52. There are typically few distinct parameter sets in a video communication session. The parameter sets are preferably transmitted reliably and out-of-band at the session set-up time. However, in some systems, mainly broadcast ones, reliable out-of-band transmission of parameter sets is not feasible, but rather parameter sets are conveyed in-band in Parameter Update NAL Packets (PUPs). Parameter sets are enumerated, and the active parameter set is indicated in the slice header.

We propose that the JVT coding standard should allow pre-definition of a large number of parameter sets and their IDs in order to avoid an excessive amount of PUPs in broadcast applications. While many of these pre-defined parameter sets are application-specific and fall outside the scope of the JVT coding standard, JVT could include a definition of default parameter sets for each profile and level.

We propose that the syntax and semantics of PUPs are defined in the normative part of the JVT standard, because PUPs are an integral part of the NAL design. Out-of-band signaling of parameter sets, such as the SDP format for parameter sets, falls out of the scope of the standard. In addition, we propose a reliable method to convey the parameter sets.

Instead of one parameter set structure, we propose four structures: independent GOP parameter set, picture parameter set, slice parameter set, and presentation parameter set. The justifications for the proposal are:

- Certain parameter values must remain unchanged throughout an independent GOP or within a picture. For example, the number of picture "slots" in the multi-picture buffer must not be changed during an independent GOP. Otherwise, it would be unclear how the multi-picture buffering process operates. A decoder implementation has to ensure that all the slices of a particular picture refer to the same picture parameter set. Similarly, all the slices of an independent GOP must refer to the same independent GOP

parameter set. Otherwise, decoders should infer data loss or corruption. Obtaining the same functionality with one joint parameter set requires that the decoder checks that individual parameter values remain the same within a picture or an independent GOP.

- A compact syntax for parameter sets is advantageous to save bits in PUPs. Thus, it makes sense to separate independent GOP and picture parameter sets from more frequently updated slice parameter sets.
- Display-related parameter values do not affect decoding of a coded video stream.

2. PRE-DEFINED PARAMETER SETS

As stated in VCEG-N52, the parameter sets are preferably transmitted reliably and out-of-band at the session set-up time. However, in some systems, mainly broadcast ones, reliable out-of-band transmission of parameter sets is not feasible, but rather parameter sets are conveyed in-band in PUPs.

However, we see the following problems:

- When there are many parameter set instances need to be transmitted in the session beginning, the out-of-band method may become overburdened or the session beginning latency will be too long.
- For systems lacking feasible mechanisms for reliable out-of-band transmission of parameter sets, in-band transport of PUPs is not reliable.
- For broadcast applications, since the parameter sets information should be transmitted frequently to allow new users join during the broadcast process, redundant transmission of all the active parameter set instances is costly from bit-rate point of view.

To tackle these problems, we propose the following method for parameter set signaling:

A relatively large number of frequently used parameter set instances and their IDs are pre-defined and stored both in the encoder and in the decoder. When the communication starts, these pre-defined parameter sets need not be transmitted. Only the parameter sets not included into the pre-defined ones have to be transmitted, preferably at the session beginning, or transmitted in PUPs later if necessary.

For applications without feedback channels (digital TV) or with feedback channels of very limited use (e.g. multicast streaming with a huge number of receivers), the set of pre-defined parameter sets should be as complete as possible, from the point of view of possible systems and application scenarios. Therefore, it may be possible that no parameter sets need to be transmitted. Once a PUP is needed, redundant transmission techniques must be applied to convey it.

One question is that in which standard the pre-defined parameter sets and their IDs are specified. We feel that the JVT standard might not be able to address the needs of variety of applications. However, some basic parameter sets, such as the default ones for each profile and level could be defined in the JVT standard.

3. PROPOSED CODING METHOD OF PARAMETER SETS

3.1 Dependencies between Coding Entities

Figure 1 presents an illustration of dependencies between NAL packets and parameter set instances. An arrow indicates a reference based on parameter set identifier. The starting point of an arrow is the entity, which refers to parameter set instance where the arrow points. The ending point of an arrow is the owner of the parameter set identifier. Single slice, data partition, and independent decoder refresh NAL packets always refer to a slice parameter set. A slice parameter set refers to a picture parameter set, and a picture parameter set refers to both an independent GOP parameter set and a presentation parameter set.

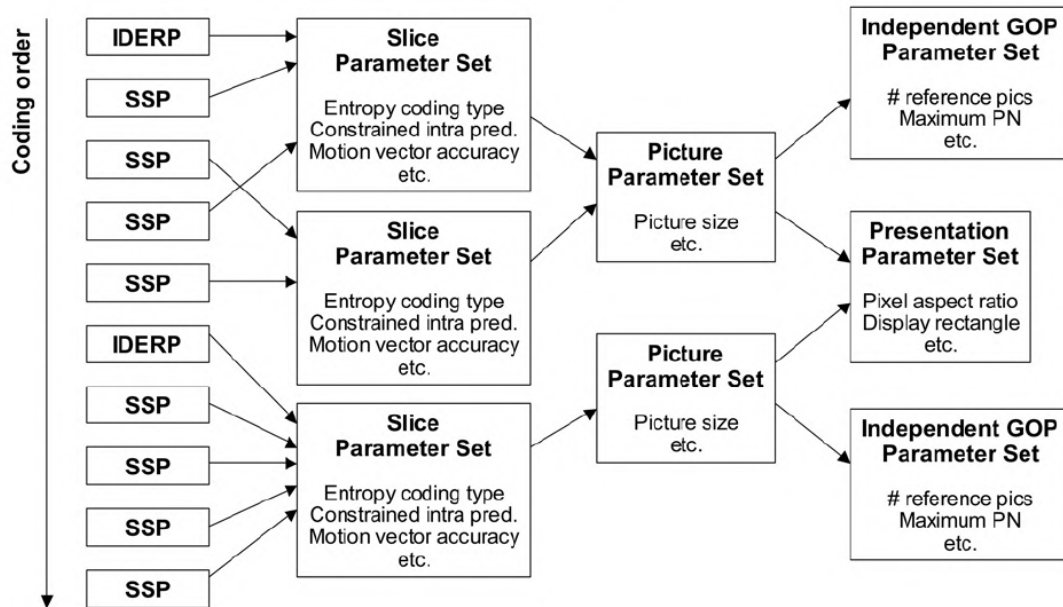


Figure 1. Illustration of dependencies between coding entities.

3.2 Notation

The notation presented in section 7.1 of JVT WD2 is used in this section to describe the syntax of the proposed parameter set structures. As we did not understand the meaning of the "Category" column in the notation, we left it blank.

3.3 On Parameter Set Identifiers

Each parameter set type has its own identifier space, i.e., a picture parameter set identifier may have the same value as a slice parameter set identifier, for example. A particular parameter set value should be unique in its identifier space.

Due to different nature of video communication applications ranging from videophones to digital TV, it is hard to define a single value for the number of allowed identifiers for all applications. For example, a conversational application may use only a couple of parameter sets, whereas some systems may pre-define a large number of parameter sets as discussed in section 2.

It can be assumed that each parameter set identifier value is referred to roughly as often as any other identifier value. Thus, variable-length coding of identifier values makes little sense. On the other hand, it would be desirable to use as few bits as possible in low-bit-rate systems using very few parameter set identifier values. As a solution, we propose two-staged coding of parameter set identifiers: First, the length of the identifier value is signaled. We propose that the length of the parameter set identifier can range from 4 to 32 bits. If necessary, the range can be stretched further, as the variable-length code-table used to signal the length is not full. Second, the identifier is fixed-length-coded using the signaled codeword length. The following pseudo-code clarifies the idea:

parameter_set_id(parameter_set_type) {	Category	Mnemonic
ec_id_length		bslbf
if (ec_id_length == 0b)		
id_length = 4		
else if (ec_id_length == 10b)		
id_length = 8		
else if (ec_id_length == 110b)		
id_length = 12		
else if (ec_id_length == 1110b)		
id_length = 16		
else if (ec_id_length == 11110b)		
id_length = 24		
else if (ec_id_length == 111110b)		
id_length = 32		
else		
reserved		
id		uimsbf(id_length)
}		

3.4 Independent GOP Parameter Set

3.4.1 Background

Independently decodable GOPs or independent GOPs were originally introduced in JVT-B041 and proposed again in JVT-C083. The start of an independent GOP is identified by an Instantaneous Decoder Refresh NAL packet (IDERP).

3.4.2 Syntax

<code>independent_GOP_parameter_set() {</code>	Category	Mnemonic
<code>parameter_set_id(independent_GOP_parameter_set)</code>		
<code>profile</code>		ecselbf
<code>level</code>		ecselbf
<code>version</code>		ecselbf
<code>log2_max_picture_number_minus_4</code>		ecselbf
<code>number_of_reference_picture_buffers_minus_1</code>		ecselbf
<code>required_picture_number_update_behavior</code>		uimsbf(1)
<code>}</code>		

3.4.3 Semantics

profile defines the coding profile in use. level defines the level in use within the profile. version defines the version in use within the profile and the level.

log2_max_picture_number_minus_4 specifies the MAX_PN constant used in picture number related arithmetic. $MAX\_PN = 2^{\log2\_max\_picture\_number\_minus\_4 + 4} - 1$. (Note: this definition assumes a fixed-length picture number field in the slice header proposed in JVT-C079. If JVT-C079 is not adopted, the definition has to be changed appropriately.)

number_of_reference_picture_buffers_minus_1 defines the total number of short- and long-term picture buffers in the multi-picture buffer.

If required_picture_number_update_behavior is 1, a specific decoder behavior in case of missing picture numbers is mandated (see section 9.1.1.2 of JVT WD2).

3.4.4 Rationale for Parameter Selection

Profile, level, and version are very likely to remain unchanged in an independent GOP.

A change in reference picture buffering controls, i.e., MAX_PN, number of reference picture buffers, and required picture number update behavior, would cause an undefined decoder state. A change of these parameters in the middle of an independent GOP would not bring any benefits. A range from 0 to 15 is considered to be a practical minimum for a picture number, and therefore smaller values are not allowed in the syntax.

3.5 Picture Parameter Set

3.5.1 Syntax

picture_parameter_set() {	Category	Mnemonic
parameter_set_id(picture_parameter_set)		
parameter_set_id(independent_GOP_parameter_set)		
parameter_set_id(presentation_parameter_set)		
picture_width_in_MBs_minus_1		ecselbf
picture_height_in_MBs_minus_1		ecselbf
}		

3.5.2 Semantics

picture_width_in_MBs_minus_1 and picture_height_in_MBs_minus_1 define the size of the picture.

3.5.3 Rationale for Parameter Selection

picture_width_in_MBs_minus_1 and picture_height_in_MBs_minus_1 could reside in the independent GOP parameter set as well, because no reference picture resampling similar to H.263 Annex P has been proposed for the JVT codec so far. However, in order to be future-proof, the parameters were located in the picture parameter set.

The presentation parameter set is referred to in picture level, as it is necessary to allow changes in the presentation parameters on picture-by-picture basis. For example, the presentation parameter set signals the display rectangle of the reconstructed pictures, which is directly related to the coded picture size signaled in the picture parameter set.

3.6 Slice Parameter Set

3.6.1 Syntax

slice_parameter_set() {	Category	Mnemonic
parameter_set_id(slice_parameter_set)		
parameter_set_id(picture_parameter_set)		
entropy_coding		ecselbf
motion_resolution		ecselbf
constrained_intra_prediction_flag		uimsbf(1)
multiple_prediction_frames_flag		uimsbf(1)
}		

3.6.2 Semantics

entropy_coding equal to zero stands for the non-arithmetic VLC coding of WD2, whereas value one stands for the arithmetic coding VLC coding of WD2.

motion_resolution equal to zero stands for 1/4-sample motion resolution, and equal to one stands for 1/8-sample motion resolution.

constrained_intra_prediction_flag equal to zero stands for normal intra prediction, whereas one stands for the constrained intra prediction. In the constrained intra prediction mode, no intra prediction is done from inter macroblocks.

multiple_prediction_frames_flag equal to zero signals that only the previous coded picture in coding order is used as a reference picture for motion compensation in P- and SP-frames and the most recent previous decoded and subsequent decoded pictures are used for motion compensation in B-frames. Otherwise, the reference picture(s) for prediction must be signaled for each macroblock.

Note: Partitioning type, which existed in earlier parameter set definitions and which signals the use of single slices or data partitions, was considered unnecessary, because NAL packet types carry the required information.

3.6.3 Rationale for Parameter Selection

The selected parameters for the slice parameter set were such that no reason to restrict their values to be unchanged in an entire independent GOP was found. Instead, it might be advantageous to allow changes in parameter values even within a picture. Examples of possible advantages include:

- Use of multiple reference pictures. It may make sense to restrict the number of reference pictures to one for certain parts of the picture. For example, in wireline video conferencing equipment, the encoder may treat the center of the picture better than the edges.
- Entropy coding type. If the isolated region concept (JVT-C072) gets approved, it is possible to mix multiple coded video streams to one without decoding the streams. This can be beneficial in a multi-point control unit (MCU) of a video conferencing system, for example. One endpoint may use a different entropy coding method than another one. Thus, the "mixed" stream would contain multiple entropy coding types in a same "mixed" coded picture.

3.7 Presentation Parameter Set

3.7.1 Syntax

presentation_parameter_set() {	Category	Mnemonic
parameter_set_id(presentation_parameter_set)		uimsbf(16)
pixel_aspect_ratio_width		uimsbf(8)
pixel_aspect_ratio_height		uimsbf(8)
display_rectangle_offset_top		uimsbf(16)
display_rectangle_offset_left		uimsbf(16)
display_rectangle_offset_bottom		uimsbf(16)
display_rectangle_offset_right		uimsbf(16)
}		

3.7.2 Semantics

The displayed pixel aspect ratio should be pixel_aspect_ratio_width:pixel_aspect_ratio_height. The parameter values shall be relatively prime. Value 0 is forbidden.

display_rectangle_offset_top, display_rectangle_offset_left, display_rectangle_offset_bottom, and display_rectangle_offset_right define the rectangle to be displayed from the coded picture. Sample units are used.

3.7.3 Default Presentation Parameter Set

We believe that a single adaptive presentation parameter set could be used for many applications. This parameter set has the following characteristics:

- Pixel aspect ratio is 1:1.
- Display rectangle is as large as the size of the reconstructed pictures.

We propose that this virtual presentation parameter set is dedicated an identifier value 0 that shall not be used when defining presentation parameter sets.

Note: Herein, the display rectangle is the part of the reconstructed image area that is to be displayed. In other words, if you decode a 176x144 picture, the encoder can signal that only the 172x140 part of it should be displayed. The display rectangle has nothing to do with the extents of the window that it is displayed on.

3.8 Parameter Update Packets

An NALP consists of an NALP header (NALPH) and an NALP payload (NALPP). The NALPH is the first byte of the NALP. The NALPH itself distinguishes the 8 defined NALP types and includes one bit indicating the presence of errors in the NALPP following the NALPH. EI flag set to 0 means that there is no known error in the following payload whereas a 1 indicates a corrupted payload and/or a corrupted NALP type.

The following table shows the NALP headers and NALP payload types. We propose signaling of the type of the targeted parameter set in the NALP header. Our syntax change proposal is highlighted in the table below.

NALPH	NALPP type	EI-Flag
0x10	SSP	0
0x11	SSP	1
0x20	DPA	0
0x21	DPA	1
0x30	DPB	0
0x31	DPB	1
0x40	DPC	0
0x41	DPC	1
0x50	SEIP	0
0x51	SEIP	1
0x60	IGOP-PUP	0
0x61	IGOP-PUP	1
0x62	PIC-PUP	0
0x63	PIC-PUP	1
0x64	SLICE-PUP	0

0x65	SLICE-PUP	1
0x66	PR-PUP	0
0x67	PR-PUP	1
0x70	CP	0
0x71	CP	1
0x80	IDERP	0
0x81	IDERP	1

The structure of the first byte, that indicates the NALP type and the status of the error indication flag, could not be chosen linearly because of design issues in other networks. In particular, this byte, preceded with a fixed two-byte start code prefix, is used as a start code in the MPEG-2 transport environment. To avoid start code emulations there, reserves the values 0x00 and 0xb9 to 0xff.

Parameter Update Packets (PUPs) contain information to update the Parameter Sets for the video stream. If the transmission and update of Parameter Sets is a function of a control protocol, PUPs should not be used in such systems where adequate protocol support is available. However, in applications where the packet stream has to be self-contained PUPs may be used.

An IGOP-PUP packet contains an independent GOP parameter set, a PIC-PUP packet contains a picture parameter set, a SLICE-PUP contains a slice parameter set, and a PR-PUP contains a presentation parameter set. The syntax for the parameter sets is presented in the previous sections of this contribution.

A PUP becomes valid synchronously with the decoding process according to the specific type of the PUP as follows: An update of an independent GOP parameter set takes place just before the first slice of the next independent GOP (i.e., the next IDERP NAL packet) is decoded. An update of picture and presentation parameter sets takes place just before the first slice of the next picture is decoded. An update of a slice parameter set takes place immediately if at least one of the contained parameters is changed. However, slice parameter set identifier used in a picture may not be redefined with different parameter values within the same coded picture.

For certain broadcast applications, such as digital TV and multicast streaming with a huge number of possible receivers, new independent GOP parameter sets, which fall out of the scope of the pre-defined ones, are likely to be updated before each independent GOP is decoded, regardless of whether or not some of the contained parameters are changed, to enable decoding in terminals that just started receiving the coded data. Other parameter sets are updated before they are referred to in the coded data.

3.9 Slice Header

The proposed slice header modifications relative to JVT WD2 are included in the accompanying annex of the contribution. In addition to including a parameter set identifier in the slice header, the annex proposes removal of redundancies and inconsistencies from the slice header. See the comments in-line for the justification of the each detailed proposal.

3.10 Examples

Our example targets at the following behavior:

- Baseline profile, level 1, version 0.

- MAX_PN = 255.
- Two reference pictures.
- Required picture number update behavior is in use.

Thus, the following independent GOP parameter set is coded and conveyed:

independent_GOP_parameter_set() {	Value	Mnemonic
parameter_set_id(independent_GOP_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
profile	0	ecselbf
level	1	ecselbf
version	0	ecselbf
log2_max_picture_number_minus_4	4	ecselbf
number_of_reference_picture_buffers_minus_1	1	ecselbf
required_picture_number_update_behavior	1	uimsbf(1)
}		

QCIF pictures are in use. Reconstructed pictures are displayed entirely, and therefore presentation parameter set 0 can be referred to. The following picture parameter set is coded and conveyed:

picture_parameter_set() {	Value	Mnemonic
parameter_set_id(picture_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
parameter_set_id(independent_GOP_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
parameter_set_id(presentation_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
picture_width_in_MBs_minus_1	10	ecselbf
picture_height_in_MBs_minus_1	8	ecselbf
}		

The encoder uses multiple reference frames for some coded pictures, while it uses a single reference picture for others. Thus, two slice parameter sets are defined as follows:

slice_parameter_set() {	Value	Mnemonic
parameter_set_id(slice_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
parameter_set_id(picture_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
entropy_coding	0	ecselbf
motion_resolution	0	ecselbf
constrained_intra_prediction_flag	1	uimsbf(1)
multiple_prediction_frames_flag	0	uimsbf(1)
}		

slice_parameter_set() {	Value	Mnemonic
parameter_set_id(slice_parameter_set) {		
ec_id_length	0b	bslbf
id	1	uimsbf(4)
}		
parameter_set_id(picture_parameter_set) {		
ec_id_length	0b	bslbf
id	0	uimsbf(4)
}		
entropy_coding	0	ecselbf
motion_resolution	0	ecselbf
constrained_intra_prediction_flag	1	uimsbf(1)
multiple_prediction_frames_flag	1	uimsbf(1)
}		

(Append for Proposal Documents)

JVT Patent Disclosure Form

International Telecommunication Union
Telecommunication Standardization Sector



International Organization for Standardization



International Electrotechnical Commission



Joint Video Coding Experts Group - Patent Disclosure Form

(Typically one per contribution and one per Standard | Recommendation)

Please send to:

JVT Rapporteur Gary Sullivan, Microsoft Corp., One Microsoft Way, Bldg. 9, Redmond WA 98052-6399, USA
Email (preferred): Gary.Sullivan@itu.int Fax: +1 425 706 7329 (+1 425 70MSFAX)

This form provides the ITU-T | ISO/IEC Joint Video Coding Experts Group (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. JVT requires that all technical contributions be accompanied with this form. *Anyone* with knowledge of any patent affecting the use of JVT work, of their own or of any other entity ("third parties"), is strongly encouraged to submit this form as well.

This information will be maintained in a "living list" by JVT during the progress of their work, on a best effort basis. If a given technical proposal is not incorporated in a Recommendation | Standard, the relevant patent information will be removed from the "living list". The intent is that the JVT experts should know in advance of any patent issues with particular proposals or techniques, so that these may be addressed well before final approval.

This is not a binding legal document; it is provided to JVT for information only, on a best effort, good faith basis. Please submit corrected or updated forms if your knowledge or situation changes.

This form is *not* a substitute for the ITU ISO IEC Patent Statement and Licensing Declaration, which should be submitted by Patent Holders to the ITU TSB Director and ISO Secretary General before final approval.

Submitting Organization or Person:

Organization name	<u>Nokia Corporation</u>
Mailing address	<u></u>
Country	<u></u>
Contact person	<u></u>
Telephone	<u></u>
Fax	<u></u>
Email	<u></u>
Place and date of submission	<u></u>

Relevant Recommendation | Standard and, if applicable, Contribution:

Name (ex: "JVT")	<u>JVT</u>
Title	<u>Coding of Parameter Sets</u>
Contribution number	<u>JVT-C078</u>

(Form continues on next page)

Disclosure information – Submitting Organization/Person (choose one box)

☒

2.0 The submitter is not aware of having any granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

or,

The submitter (Patent Holder) has granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution. In which case,

☐

2.1 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a free license to an unrestricted number of applicants on a worldwide, non-discriminatory basis to manufacture, use and/or sell implementations of the above Recommendation | Standard.

☐

2.2 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and on reasonable terms and conditions to manufacture, use and/ or sell implementations of the above Recommendation | Standard.

Such negotiations are left to the parties concerned and are performed outside the ITU | ISO/IEC.

☐

2.2.1 The same as box 2.2 above, but in addition the Patent Holder is prepared to grant a “royalty-free” license to anyone on condition that all other patent holders do the same.

☐

2.3 The Patent Holder is unwilling to grant licenses according to the provisions of either 2.1, 2.2, or 2.2.1 above. In this case, the following information must be provided as part of this declaration:

- patent registration/application number;
- an indication of which portions of the Recommendation | Standard are affected.
- a description of the patent claims covering the Recommendation | Standard;

*In the case of any box **other than 2.0** above, please provide the following:*

Patent number(s)/status _____

Inventor(s)/Assignee(s) _____

Relevance to JVT _____

Any other remarks: _____

(please provide attachments if more space is needed)

(form continues on next page)

Third party patent information – fill in based on your best knowledge of relevant patents granted, pending, or planned by other people or by organizations other than your own.

Disclosure information – Third Party Patents (choose one box)

☒

3.1 The submitter is not aware of any granted, pending, or planned patents *held by third parties* associated with the technical content of the Recommendation | Standard or Contribution.

☐

3.2 The submitter believes third parties may have granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

For box 3.2, please provide as much information as is known (provide attachments if more space needed) - JVT will attempt to contact third parties to obtain more information:

3rd party name(s)

Mailing address

Country

Contact person

Telephone

Fax

Email

Patent number/status

Inventor/Assignee

Relevance to JVT

Any other comments or remarks: