

EXHIBIT 2004

Title: **Absolute signaling of reference pictures**

Status: Input Document to JCT-VC

Purpose: Proposal

Author(s) or Rickard Sjöberg, Jonatan Samuelsson

Contact(s): Ericsson Research
164 80 Stockholm
Sweden

Tel: +46 761153563

Email: rickard.sjoberg@ericsson.com
jonatan.samuelsson@ericsson.com

Source: Ericsson

Abstract

This document contains a proposal regarding the reference picture processes in HEVC. It is proposed that pictures that should be available for reference are described in an absolute manner in the slice header of a picture instead of in a relative manner (through MMCO and sliding window process). It is further proposed that buffer operations (picture marking) is performed directly after parsing of the first slice header of a picture instead of after the picture has been decoded.

The document claims that the proposed changes improves error resilience, simplifies temporal scalability and enables decoding at a minimum number of decoded picture buffer slots.

1 Current decoded picture buffer scheme

The processes for reference picture handling and decoded picture buffer operations in the working draft of HEVC [1] are to a large extent inherited from H.264/AVC although some sections are still missing in the working draft text. A simplified flow chart of the processes in H.264/AVC is shown in *Figure 1*.

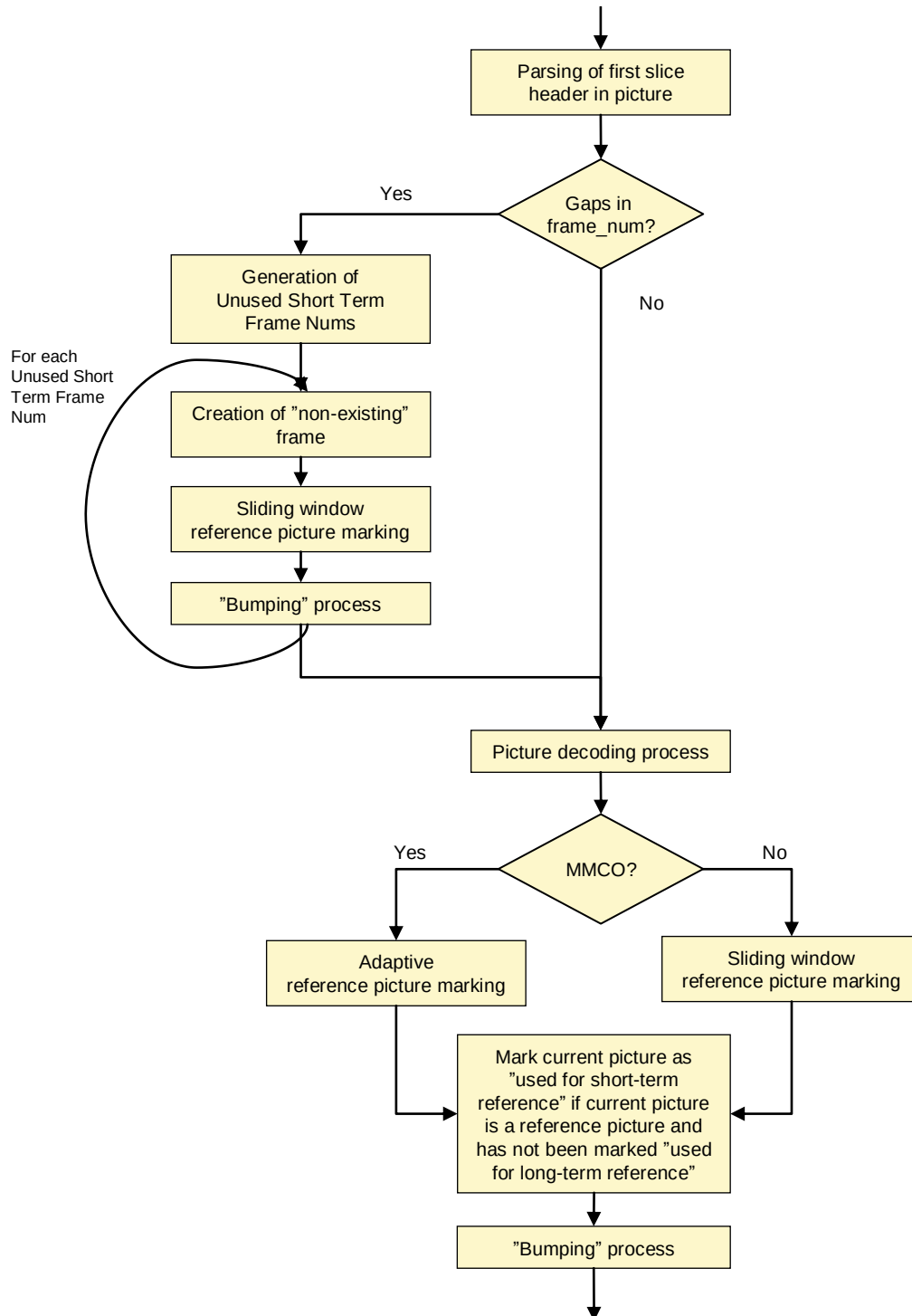


Figure 1. Current HEVC scheme.

Before the actual decoding of a picture, the frame_num in the slice header is parsed to detect a possible gap in frame_num. If there is a gap in frame_num, “non-existing” frames are created and inserted into the Decoded Picture Buffer (DPB) using the sliding window process to maintain the correct number of reference pictures and the “bumping” process to deliver decoded pictures in correct order.

Regardless of whether there was a gap in frame_num or not the next step is the actual decoding of the current picture. If the slice headers of the picture contain MMCO commands, adaptive memory control process is applied after decoding of the picture; otherwise the sliding window process is applied. As a final step, the “bumping” process is applied.

1.1 Problems with existing scheme

1.1.1 Error resilience

A problem with H.264/AVC is its vulnerability to losses of pictures that contains Memory Management Control Operations (MMCO) of type 2, 3, 4, 5 or 6 (as described in Table 1 copied from H.264/AVC specification).

memory_management_control_operation	Memory Management Control Operation
0	End memory_management_control_operation syntax element loop
1	Mark a short-term reference picture as "unused for reference"
2	Mark a long-term reference picture as "unused for reference"
3	Mark a short-term reference picture as "used for long-term reference" and assign a long-term frame index to it
4	Specify the maximum long-term frame index and mark all long-term reference pictures having long-term frame indices greater than the maximum value as "unused for reference"
5	Mark all reference pictures as "unused for reference" and set the MaxLongTermFrameIdx variable to "no long-term frame indices"
6	Mark the current picture as "used for long-term reference" and assign a long-term frame index to it

Table 1. Memory management control operation values from H.264/AVC.

Loss of a picture that does not contain MMCO, or a picture that contains MMCO of type 1, is of course severe to the decoding process. Pixel values of the lost picture will not be available and may affect future picture for a long period of time due to incorrect inter prediction. There is also a risk that reference picture lists for a few pictures following the lost picture will be wrong, for example if the lost picture contained MMCO that marked one short-term reference picture as “unused for reference” that otherwise would have been included in the reference picture list of the following picture. However, the decoding process can recover such a loss through usage of intra slices, constrained intra blocks or by other means.

But if a picture containing MMCO of type 2, 3, 4, 5 or 6 is lost there is a risk that the number of long term pictures in the DPB is different from what it would have been if the picture was received, resulting in an “incorrect” sliding window process for all the following pictures. That is, the encoder and decoder will contain a different number of short-term pictures resulting in out-of-sync behavior of the sliding window

process. This loss cannot be recovered through usage of constrained intra blocks, intra slices or similar techniques (not even an open GOP Intra picture). The only way to ensure recovery from such a loss is through an IDR picture or through an MMCO that cancels the effect of the lost MMCO. What makes the situation even worse is that a decoder will not necessarily know that the sliding window process is out-of-sync and thus cannot report the problem to the encoder or request an IDR picture even in applications where a feedback channel is available.

One way to reduce the risk of losing important MMCO information is to use dec_ref_pic_marking_repetition SEI messages. However the encoder will (most probably) not know if the decoder is capable of making use of dec_ref_pic_marking_repetition SEI messages. Further, there is a risk that the dec_ref_pic_marking_repetition SEI message is also lost.

1.1.2 Temporal scalability

Temporal scalability is supported in H.264/AVC (and SVC) through the definition of sub-sequences (and usage of temporal_id in SVC) and insertion of “non-existing” frames. However, in order to support temporal scalability, the pictures in the higher temporal layers are restricted when it comes to usage of MMCO. The encoder is responsible of making sure that the MMCO commands in one temporal layer does not affect pictures of lower temporal layers differently compared to if the temporal layer is dropped and “non-existing” pictures are inserted and sliding window process is applied.

This imposes restrictions on the encoder in its selection of coding structure and reference picture usage. Consider the example in **Figure 2** which is a low-delay configuration with two temporal layers using two reference pictures for each picture.

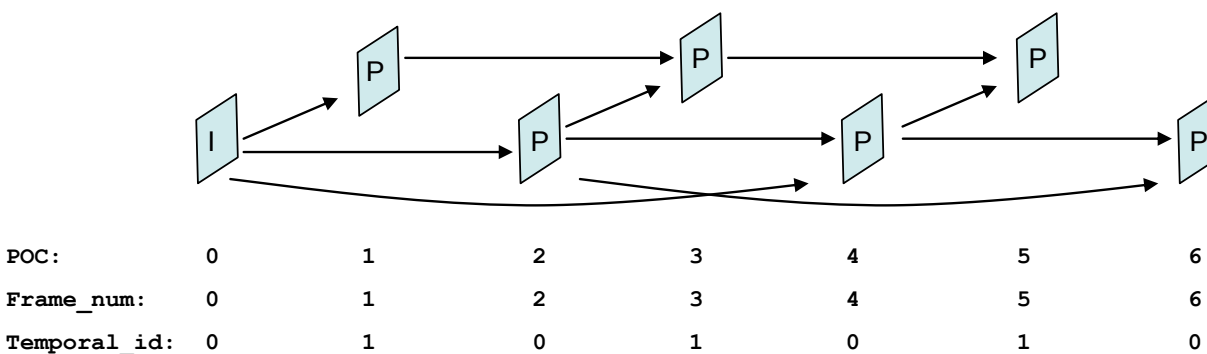


Figure 2. Example low delay coding structure with two temporal layers.

The minimum value of max_num_ref_frames for this coding structure is 3. Each picture only uses 2 reference pictures for inter prediction but each picture must hold one extra picture from the other temporal layer that will be used for prediction by the next picture.

In order to have picture 0 and picture 2 available when decoding picture 4, picture 3 must have an MMCO command marking picture 1 as unavailable.

However, if temporal layer 1 is removed (for example by a network node) there will be gaps in frame_num for all odd numbered pictures. “Non-existing” pictures will be created for these pictures and sliding window process will be applied. That will result in having the “non-existing” picture 3 marking picture 0 as unavailable. Thus picture 0 will not be available for prediction when picture 4 is decoded. Since the encoder cannot make the decoding process be the same for the two cases; when all picture are decoded and when only the lowest layer is decoded, the example coding structure cannot be used for temporal scalability given a max_num_ref_frames of 3. This problem is caused by MMCO commands in a higher layer when later this higher layer is removed since then sliding window is assumed for the missing pictures.

It is explicitly specified in the SVC specification text that:

“With currTId being the value of temporal_id for the current access unit, the bitstream shall not contain data that result in the marking of pictures with temporal_id less currTId as "unused for reference" during this invocation of the process in subclause G.8.2.4.4.” Where G.8.2.4.4. is the “SVC adaptive memory control decoded reference picture marking process” section.

1.1.3 Gaps in frame_num and B slices reference picture list initialization

1.1.3.1 Missing POC values

Reference picture list initialization for B slices uses the POC value of the pictures in the DPB in order to create the correct reference picture lists.

If POC type 0 is used in H.264/AVC, each picture must explicitly signal its POC. A “non-existing” picture that is inserted into the DPB due to a gap in frame_num is not associated with a POC when POC type 0 is used.

To ensure correct decoding process, for every B picture that is received when one of the reference pictures is a “non-existing” picture, the reference picture list modification will have to be used, to explicitly signal the reference pictures in every slice of the picture. This is mentioned in a note in the H.264/AVC specification:

“NOTE 1 – When gaps_in_frame_num_value_allowed_flag is equal to 1, encoders should use reference picture list modification to ensure proper operation of the decoding process (particularly when pic_order_cnt_type is equal to 0, in which case PicOrderCnt() is not inferred for "non-existing" frames).”

The signaling of reference picture list modifications will increase the bit overhead of each slice header.

1.1.3.2 Missing temporal_id values

In the current HEVC working draft, reference picture list initialization for B slices uses the temporal_id value of the pictures in the DPB in order to create the correct reference picture lists. A picture in the DPB with higher temporal_id than the current picture is not included in the reference picture list for the current picture.

The current working draft of HEVC does not specify constraints on temporal_id or gaps in frame num nor methods for how to assign temporal_id to “non-existing” pictures. If a “non-existing” picture is not given a temporal_id value it is not possible for a decoder to know if the “non-existing” picture should be included or not during reference picture list initialization.

2 Discussion

What we would like to see is a scheme that improves error robustness compared to H.264/AVC by ensuring that the loss of a picture does not affect the reference picture lists for following pictures other than that the pixel values and motion vectors of the lost picture will not be available.

We would also like to see a scheme that is flexible and well suited for temporal scalability.

We would also like to see a scheme that minimizes overhead of picture buffer operations and reference picture list initializations and we would like it to be straight forward to use a minimum number of reference picture slots in the decoded picture buffer.

3 Proposal

These are the changes that we propose:

- Move decoded reference picture marking from after decoding a picture to after parsing of the first slice header of a picture
- Add explicit signaling of the reference pictures, using a Buffer Description, in each slice header, containing information to deduce POC and temporal_id of all reference pictures marked as “used for reference” in the DPB
- Introduce the possibility to include a list of Buffer Descriptions in the Picture Parameter Set to minimize the overhead of signaling available reference pictures (Buffer Descriptions) in the slice header
- Remove all MMCO commands as the functionality provided by them is covered by Buffer Descriptions
- Remove the sliding window operation as this functionality is provided by Buffer Descriptions
- Require explicit signaling of POC (i.e. remove pic_order_cnt_type 1 and 2 as well as the syntax element pic_order_cnt_type)
- Use POC for identification of pictures during reference picture list construction and decoded reference picture marking
- Remove the syntax elements frame_num
- Remove the syntax element gaps_in_frame_num_allowed_flag
- Remove long term picture marking as this functionality is provided by Buffer Descriptions
- Remove “non-existing” frames
- Introduce the rule that a Buffer Description must not contain a picture B if the temporal_id of B is less than or equal to the temporal_id of the current picture unless that picture B is available in the decoded picture buffer and marked as “used for reference”.

These changes are explained in details in the rest of this document.

We propose a scheme where all picture marking operations are applied after parsing of the first slice header of a picture but before reference picture list initialization and before picture decoding, using a description of the reference pictures, a “Buffer Description” as illustrated in **Figure 3**. The Buffer Description is signaled in the slice header either explicitly or by reference to a predefined structure (Buffer Description) signaled in a Picture Parameter Set.

It can be noted that the proposal implies conceptual changes to the decoding process. In H.264/AVC and current design of HEVC, operations are given to the decoder either implicitly (sliding window) or explicitly (MMCO) and the decoder is responsible of applying these operations and keeping track of the reference pictures (i.e. which pictures can be used for reference). In the proposed scheme the reference pictures (i.e. which pictures can be used for reference) is signaled within the current picture (in the slice header) thus removing the need of implicitly and explicitly signaled operations.

This means that each picture will have an absolute description of the reference pictures instead of a relative description as in H.264/AVC (where delta-information is retrieved from MMCO or from the standard text, i.e. sliding window process).

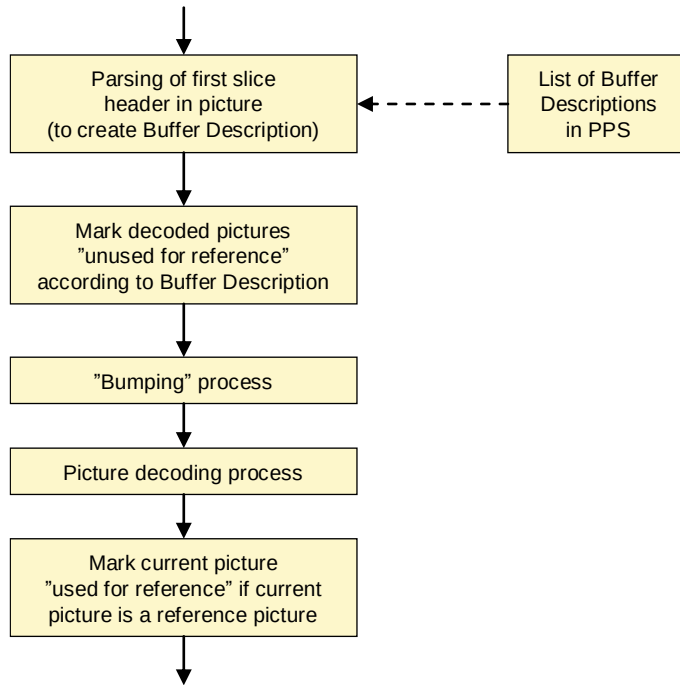


Figure 3. Proposed scheme.

3.1 Buffer Description

It is proposed that a Buffer Description contains *deltaPOC* information and *temporal_id* of all reference pictures that the decoder shall keep. The *deltaPOC* is used to calculate the POC of a reference picture as $POC_{reference} = POC_{current} + \text{deltaPOC}$. POC is used by the decoder to deliver the pictures in the correct order for display but also for identification of reference pictures during reference picture list construction and decoded reference picture marking. *temporal_id* is included in the description to enable correct reference picture list construction in the case of lost or removed pictures (e.g. removed by temporal scalability).

All pictures in the DPB that are not a part of the Buffer Description shall be marked as “unused for prediction”. Note that they may still be kept in the decoded picture buffer for the bumping process.

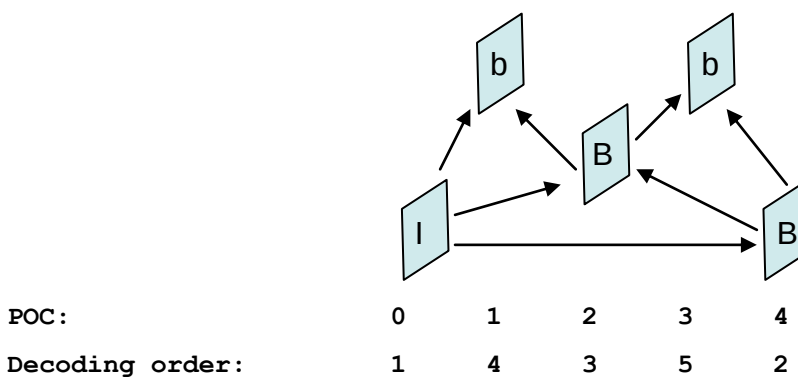


Figure 4. Example coding structure.

In the example in **Figure 4**, the second picture in decoding order will include POC 0 in the Buffer Description while the third picture in decoding order will include POC 0 and POC 4 in its Buffer Description. The fourth picture in decoding order must include not only POC 0 and POC 2 in its Buffer Description but also POC 4 since that picture will be used for reference in the future. The fifth picture in decoding order does not need to include POC 0 in its Buffer Description unless it is going to be used for

reference in the future. If POC 0 is not included in the Buffer Description of the fifth picture in decoding order it will be marked “unused for prediction”.

3.1.1 A list of Buffer Descriptions in the Picture Parameter Set

To reduce the bit overhead of signaling the Buffer Descriptions explicitly in the slice header it is proposed to introduce the possibility to send a list of Buffer Descriptions in the Picture Parameter Set. These Buffer Descriptions can be used for multiple slices and multiple pictures in the same sequence and thus significantly reduce the number of bits required per picture.

For each Buffer Description in the Picture Parameter Set the deltaPOC information and temporal_id of all reference pictures should be present.

3.1.2 Removal of frame_num and gaps_in_frame_num_allowed_flag

In H.264/AVC the syntax element frame_num is used to identify pictures in the DPB and to detect gaps in frame num. If gaps_in_frame_num_allowed_flag is equal to 1 the decoder shall insert “non-existing” frames in the DPB in order for the sliding window process to operate correctly.

In the proposed scheme the POC is used to identify pictures in the DPB. The proposed scheme does not contain a sliding window process. Therefore, it is proposed to remove the syntax elements frame_num and gaps_in_frame_num_allowed_flag.

3.1.3 Removal of “non-existing” frames and a bitstream requirement to formalize temporal scalability and enable detection of lost reference pictures

In H.264/AVC a gap in frame num invokes a process for creating “non-existing” frames. We propose to remove the syntax elements frame_num and gaps_in_frame_num_allowed_flag as described above.

It is obvious that a gap in frame_num cannot be detected if the syntax element frame_num is removed. We propose that the generation of “non-existing” frames is removed altogether.

Instead we propose to add a requirement on the bitstream regarding Buffer Descriptions that makes sure that the decoding process will be correct regardless of how many temporal layers are decoded and that solves the problems described in section 1.1.3.

The requirement on the bitstream is that no picture A encoded in the bitstream shall use a Buffer Description that contains a picture B with temporal_id lower than or equal to the temporal_id of A unless B is available in the DPB and marked as “used for prediction”.

This requirement together with the rule that no pictures with temporal_id higher than the current temporal_id are put in the reference picture lists ensures that the reference picture lists will be correct in the decoder without explicit signaling of reference picture list modifications regardless of how many temporal layers are decoded. The requirement also makes ensures that generation of “non-existing” frames is not needed.

A bitstream that includes a picture A encoded in the bitstream that uses a Buffer Description that contains a picture B with temporal_id lower than or equal to the temporal_id of A where B is not available in the DPB or marked as “unused for prediction” should be interpreted by the decoder as a bitstream that has been subject to an unintentional picture loss and B should be inferred to be an unintentionally lost picture. This means that unintentional picture losses of reference pictures are detected when parsing the slice header.

If the picture B that is included in Buffer Description of picture A has a higher temporal id than the temporal id of A no action is performed. It does not matter for the decoding process of A whether B is available in the DPB or not since B is not included in the reference picture lists of A. If B is removed due to temporal scaling the decoding process will continue without being affected. If B is unintentionally lost

the loss will be detected as soon as a picture with temporal_id higher than or equal to the temporal_id of B contains B in its Buffer Description.

With this requirement we remove the possibility of sub-sequences within the same temporal layer. Sub-sequences have to use temporal_id layering to not break the requirement above. Note that this does not remove any of the functionalities provided by sub-sequences.

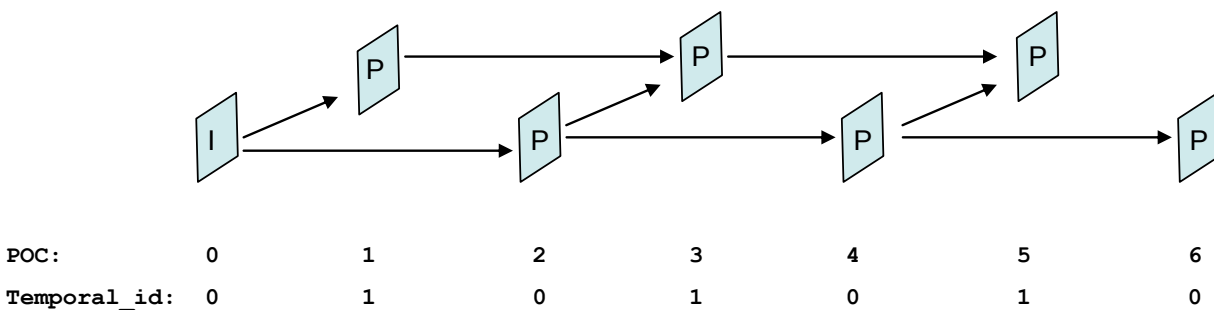


Figure 5. An example of temporal scaling.

Consider the example illustrated in figure 5 where POC numbers are used to identify pictures. Picture 4 must have picture 3 in its buffer description otherwise picture 3 would be marked as “unused for reference” and could not be used by picture 5. But if only the lowest temporal layer is decoded, picture 3 is not available in the DPB when picture 4 is decoded, but that is not a problem since picture 4 can not use picture 3 for prediction nor can any other picture in the lowest temporal layer.

Picture 4 also contains picture 2 in its buffer description. With the proposed bitstream requirement, picture 2 can not be removed since it is in the same temporal layer as picture 4. If picture 2 is not available in the DPB and marked “used for reference” the decoder can instantly draw the conclusion that picture 2 has been lost since the encoder or a network node can not remove picture 2 due to the proposed bitstream requirement.

Thus, losses of reference pictures are always detected which is very useful for a decoder in order to take proper action by for example requesting refresh. Also, the POC of a lost reference picture is always known which can be very useful for error concealment.

3.2 Rules for Buffer Descriptions

An IDR picture shall not send a Buffer Description since it does not use any previous pictures for reference and removes all reference pictures from the DPB.

For all non-IDR pictures a Buffer Description is **required** in order to apply reference picture marking before reference picture list initialization and decoding of the current picture.

A Buffer Description may have at most max_num_ref_frames pictures in its description.

The Buffer Description is signaled either explicitly or via reference to a Buffer Description in a Picture Parameter Set.

Reference pictures that are not included in the Buffer Description will be marked “unused for reference”.

A picture that is used for reference for a long period of time (long term reference picture) is preferably signalled explicitly in the slice headers of the pictures for which it is available for reference, since it would be inconvenient to include every distance (in POC numbers) in the Buffer Description in the PPS.

All pictures in a Buffer Description must have temporal_id set to the temporal_id of the corresponding reference picture.

3.3 Graphical illustration of Buffer Description application example with $\text{max_num_ref_frames} = 3$

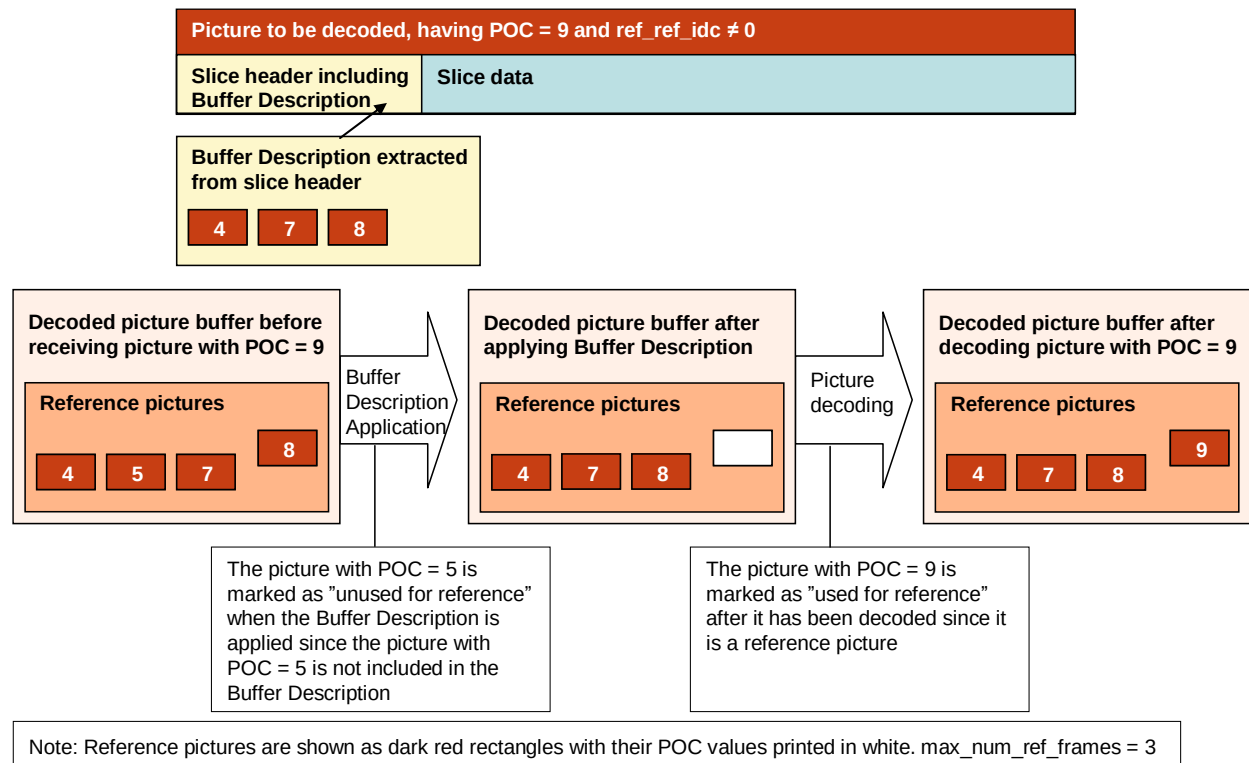


Figure 6. A graphical illustration of how a Buffer Description is applied

3.4 Implications on display of pictures

In H.264/AVC the process that delivers pictures for output (referred to as “bumping” process in figure 1) is sometimes performed prior to decoding (if there was a gap in frame num). The “bumping” process is also performed after decoding (and picture marking).

In the proposed scheme the “bumping” process is always applied prior to decoding. It could be argued that this imposes extra delay in the decoding process before delivery of pictures for output. However it should be noted that the first picture to display is uniquely defined already after the decoding process step as soon as the number of non-displayed pictures in DPB is larger than or equal to $\text{num_reorder_frames}$. Thus a decoder can deliver that picture for display directly after the decoding process step. It would be possible to add such a step to the scheme as illustrated in **Figure 7**. Alternatively, it could be mentioned in a note in the standard text that this action is possible to reduce delay. Thus the delay of the proposed scheme is equal to the delay of the current HEVC scheme.

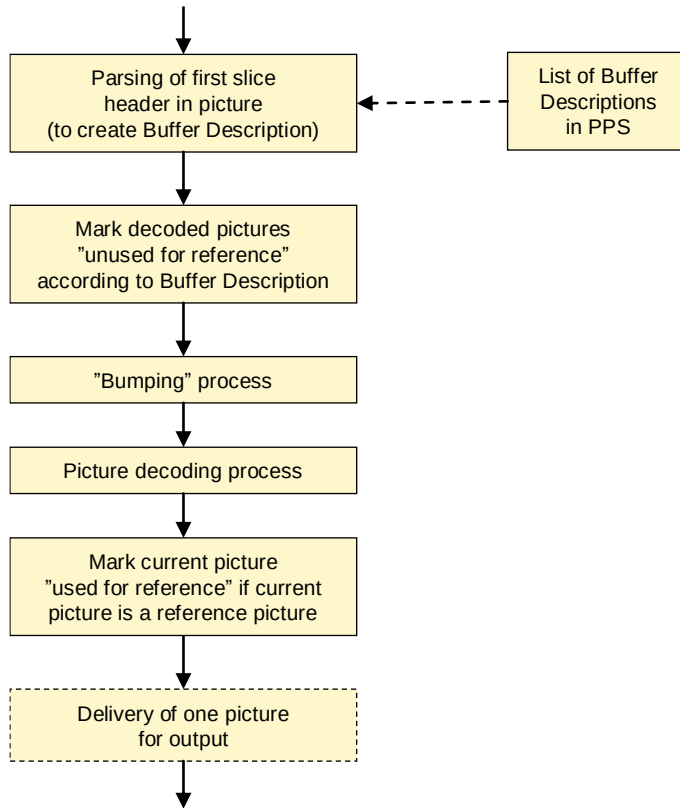


Figure 7. Alternative scheme.

3.5 Advantages of the proposal

The proposed solution increases the error robustness level of HEVC. Instead of relying on correctly delivered and interpreted buffer operations in previous pictures, the decoder will only have to rely on information contained in the slice header(s) of the current picture. As the decoded picture marking operations are based on the Buffer Description of the current picture and applied before decoding of the current picture there is no longer a risk that a lost picture affects the reference picture lists of succeeding pictures.

With the proposed scheme a decoder will always know when picture that is used for reference has been lost. The decoder will also know the POC value of a lost reference picture, which is very useful for error concealment.

Using the proposed solution it is possible to apply the optimal buffer scheme using as few image buffers as possible for any coding structure.

The proposed solution enables a consistent scheme for temporal scalability that is easy for an encoder and a network node to exploit. The proposed solution provides good error robustness also for temporally scaled bitstreams.

The proposed solution reduces the complexity and logic required in both encoder and decoder as it replaces MMCO, sliding window, gaps in frame num, "non-existing" pictures and long term pictures with one compact solution.

The signaling of Buffer Descriptions in the PPS can decrease the bitrate for many coding structures compared to using MMCO and reference picture list modifications, yet it provides added robustness.

3.6 Working draft text

Proposed changes to the Working Draft are attached in a separate document. The proposed working draft text puts the list of Buffer Descriptions in the PPS. We propose to put the list syntax in the SPS if JCT-VC decides to put ALF syntax or similar in the PPS which then would result in one PPS per picture in the bitstream. Otherwise the bit cost would be unnecessarily high. The proposed working draft text puts Buffer Description signaling in the slice header. If a slice parameter set or any other picture header type of syntax is adopted by JCT-VC, we propose to move the Buffer Description signaling from the slice header to such a slice parameter set or picture header since there is no need to signal it repeatedly in every individual slice header in one picture.

An addition to the proposed syntax in the attached Working Draft changes is to allow for both referenced and explicit coding of a Buffer Description. First there is a flag whether reference or explicit Buffer Description is used. If explicit signaling is signaled, there is a second flag whether combined coding is used or not. If combined coding is off, there is only explicit coding of the Buffer Description. If combined coding is on, there is both reference and explicit coding. First the referenced Buffer Description is decoded and then a $ue(v)$ codeword is present to indicate how many reference pictures that in addition are explicitly signaled. The explicitly signaled reference pictures replace entries in the Buffer Description starting from the end. The explicitly signaled reference pictures in the combination are signaled the same way as the pure explicitly signaled reference pictures with the exception that the POC for the explicitly signaled reference pictures in the combination are coded with $\log_2_max_pic_order_cnt_minus4+4$ bits instead of $ue(v)$, where $\log_2_max_pic_order_cnt_minus4$ is a $ue(v)$ syntax element in the SPS that describes the number of bits to use for POC.

3.7 Example list of Buffer Descriptions

Figure 8 shows an example low-delay coding structures with three temporal layers, similar to the low delay test case in common test conditions.

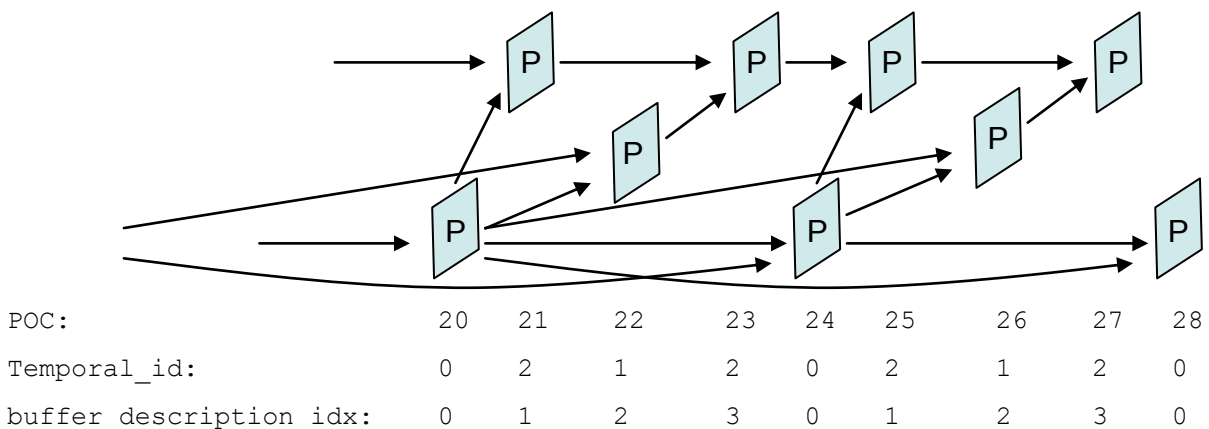


Figure 8. Coding structure with (sub)GOP size 4. In the figure, the usage of two reference pictures are illustrated. However, a total of at least 4 reference pictures must be used to enable the illustrated coding structure.

An example list of Buffer Descriptions for the coding structure in **Figure 8** can be seen in Table 2 where 4 reference pictures are used and where deltaPOC (dP) and temporal_id (tId) are indicated for each picture in each Buffer Description. The list of Buffer Description is constructed using the scheme of including the two closest reference pictures (POC(current)-1 and POC(current)-2) and the two from the lowest temporal layer in the range from POC(current)-3 to POC(current)-10.

Buffer Description Index	Picture buffer 1	Picture buffer 2	Picture buffer 3	Picture buffer 4
0	dP: -1 tId: 2	dP: -2 tId: 1	dP: -4 tId: 0	dP: -8 tId: 0
1	dP: -1 tId: 0	dP: -2 tId: 2	dP: -5 tId: 0	dP: -9 tId: 0
2	dP: -1 tId: 2	dP: -2 tId: 0	dP: -6 tId: 0	dP: -10 tId: 0
3	dP: -1 tId: 1	dP: -2 tId: 2	dP: -3 tId: 0	dP: -7 tId: 0

Table 2. Example Buffer Description List. dP is for deltaPOC and tId for temporal_id.

The encoder can signal the list of Buffer Descriptions shown in Table 2 in a PPS and then signal buffer_description_idx for each picture as indicated in **Figure 8**.

The number of bits required to signal the Buffer Description List for subGOP size 4 according to Table 2 using the proposed syntax is 117. The cost for signaling buffer_description_reference_flag and buffer_description_id in the slice header is for this case on average 4 bits. Assuming the syntax element frame_num, that is proposed to be removed, is encoded with 8 bits, the proposed scheme reduces the overhead compared to the current HEVC scheme as soon as the sequence contains at least 30 slices.

4 Source code

The proposed scheme for reference picture buffer handling has been implemented in HM 3.1.

The complete source code is provided as an attachment to this contribution.

Two new config file parameters have been added; MaxNumberOfReorderPictures and MaxNumberOfReferencePictures.

5 Results

The common test conditions [2] have been run using Buffer Descriptions signaled in PPS. The order in which pictures are processed has not been changed nor which frames are used for reference. Results are provided in an attached excel sheet and a summary is provided in Table 3. The complexity numbers come from cluster simulations using different machines and are therefore not reliable.

For the Random Access coding condition HEVC requires allocation of 8 reference pictures in the decoder for sliding window to operate correctly without MMCO. But it can be noted that HM 3.1 allocates 12 pictures in the decoder. This is probably due to that frame_num is missing in the code as well as the sliding window process.

Our implementation for the Random Access coding condition allocates 6 reference pictures in the decoder which is the optimal number for that coding structure and picture coding order.

	All Intra HE			All Intra LC		
	Y	U	V	Y	U	V
Class A	0,0	0,0	0,0	0,0	0,0	0,0
Class B	0,0	0,0	0,0	0,0	0,0	0,0
Class C	0,0	0,0	0,0	0,0	0,0	0,0
Class D	0,0	0,0	0,0	0,0	0,0	0,0
Class E	0,0	0,0	0,0	0,0	0,0	0,0
Overall	0,0	0,0	0,0	0,0	0,0	0,0
Enc Time[%]	101%			101%		
Dec Time[%]	103%			99%		
	Random Access HE			Random Access LC		

	Y	U	V	Y	U	V
Class A	0,0	0,0	0,0	0,0	0,0	0,0
Class B	0,0	0,0	0,0	0,0	0,0	0,0
Class C	0,0	0,0	0,0	0,0	0,0	0,0
Class D	0,1	0,1	0,1	0,1	0,1	0,1
Class E						
Overall	0,0	0,0	0,0	0,0	0,0	0,0
Enc Time[%]	100%			101%		
Dec Time[%]	106%			101%		
	Low delay B HE			Low delay B LC		
	Y	U	V	Y	U	V
Class A						
Class B	0,0	0,0	0,0	0,0	0,0	0,0
Class C	0,0	0,0	0,0	0,0	0,0	0,0
Class D	0,0	0,0	0,0	0,0	0,0	0,0
Class E	0,0	0,0	0,0	0,0	0,0	0,0
Overall	0,0	0,0	0,0	0,0	0,0	0,0
Enc Time[%]	100%			101%		
Dec Time[%]	103%			102%		

Table 3. Results for common test conditions.

Some minor increases in BD-rate can be seen in Table 3. It can be noted that the anchor does not contain the syntax element frame_num described by the current Working Draft of HEVC. If an 8 bit frame_num is added to each slice header of the anchor the proposed method gives instead minor gains in BD-rate compared to that anchor as shown in an attached excel sheet.

With the proposed scheme for Buffer Descriptions it is very easy to try different structures for reference picture usage. The only thing needed to be changed is the list of Buffer Descriptions encoded in the PPS.

One scheme that has been evaluated for the Low Delay coding conditions is when the two closest pictures are used for reference plus the two pictures with lowest QP in the range from POC(current)-3 to POC(current)-10. That is the same coding structure as described by the example Buffer Description in Table 2 but with all temporal_ids set to 0.

The decoder is not modified at all to handle this case and the number of allocated reference pictures is still 4. The results are attached and a summary can be seen in Table 4.

	Low delay B HE			Low delay B LC		
	Y	U	V	Y	U	V
Class A						
Class B	-1,8	-2,8	-2,8	-2,1	-3,5	-4,4
Class C	-3,1	-3,3	-3,5	-3,2	-4,0	-4,1
Class D	-3,3	-3,0	-3,3	-3,7	-5,7	-5,7
Class E	-0,6	-0,5	-0,9	-1,1	-1,4	-1,9
Overall	-2,2	-2,6	-2,7	-2,6	-3,8	-4,2
Enc Time[%]	104%			104%		
Dec Time[%]	103%			103%		

Table 4. Results for LD with two low QP reference frames.

In the attached source code this structure is turned on by defining ERICSSON_USE_SPECIALIZED_BD_FOR_LD to 1 in CommonDef.h. If it is defined to be 0 the structure of the common test conditions is used.

6 References

- [1] JCTVC-E603, Working Draft 3 of High-Efficiency Video Coding Associated Resources, T. Wiegand, B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan.
- [2] JCTVC-E700, Common test conditions and software reference configurations, F. Bossen

7 Patent rights declaration(s)

Ericsson may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).