

EXHIBIT 2008

**DOCUMENT MADE AVAILABLE UNDER THE
PATENT COOPERATION TREATY (PCT)**

International application number:	PCT/SE2012/050439
International filing date:	26 April 2012 (26.04.2012)
Document type:	Certified copy of priority document
Document details:	Country/Office: US
	Number: 61/503,019
	Filing date: 30 June 2011 (30.06.2011)
Date of receipt at the International Bureau:	05 May 2012 (05.05.2012)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a),(b) or (b-bis)

30-04-2012

PA 7344511



THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 09, 2012

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A
FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 61/503,019

FILING DATE: June 30, 2011

THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY
APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS
CONVENTION, IS US61/503,019

By Authority of the

Under Secretary of Commerce for Intellectual Property
and Director of the United States Patent and Trademark Office



N. Williams
N. WILLIAMS
Certifying Officer

Electronic Acknowledgement Receipt	
EFS ID:	10425370
Application Number:	61503019
International Application Number:	
Confirmation Number:	8712
Title of Invention:	ABSOLUTE REFERENCE PICTURE SIGNALING
First Named Inventor/Applicant Name:	Jonatan Samuelsson
Customer Number:	27045
Filer:	Steven Ware Smith/Jennifer Hardin
Filer Authorized By:	Steven Ware Smith
Attorney Docket Number:	
Receipt Date:	30-JUN-2011
Filing Date:	
Time Stamp:	13:03:17
Application Type:	Provisional

Payment information:

Submitted with Payment	yes
Payment Type	Deposit Account
Payment was successfully received in RAM	\$220
RAM confirmation Number	11168
Deposit Account	501379
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows: Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees) Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)	

Charge any Additional Fees required under 37 C.F.R. Section 1.19 (Document supply fees)
 Charge any Additional Fees required under 37 C.F.R. Section 1.20 (Post Issuance fees)
 Charge any Additional Fees required under 37 C.F.R. Section 1.21 (Miscellaneous fees and charges)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		P34680-US1_2011-06-30_11-3698_Provisional_Application_Transmittal_Letter.pdf	90474 376ff522ced359bf3e40e0a7712c5adba1d8ff	yes	4
Multipart Description/PDF files in .zip description					
	Document Description	Start	End		
	Provisional Cover Sheet (SB16)	1	2		
	Authorization to access Appl. by Trilateral Office	3	4		

Warnings:

Information:

2	Specification	P34860-US1_2011-06-30_11-3698_Specification.pdf	145422 44033c4f58363ea5be1594870f517d6b33162001	no	21
---	---------------	---	--	----	----

Warnings:

Information:

3	Fee Worksheet (SB06)	fee-info.pdf	29135 3cb815c2825018747b74783b7e370dea99751a84	no	2
---	----------------------	--------------	---	----	---

Warnings:

Information:

Total Files Size (in bytes):	265031
-------------------------------------	--------

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. Transmitted Via EFS-Web

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Jonatan	Samuelsson	Stockholm, Sweden
Rickard	Sjoberg	Stockholm, Sweden

Additional inventors are being named on the _____ separately numbered sheets attached hereto

TITLE OF THE INVENTION (500 characters max):

Absolute Reference Picture Signaling

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☒ The address corresponding to Customer Number: 27045

OR

☐ Firm or Individual Name _____

Address _____

City	State	Zip
Country	Telephone	Email

ENCLOSED APPLICATION PARTS (check all that apply)

☐ Application Data Sheet. See 37 CFR 1.76 ☐ CD(s), Number of CDs _____

☐ Drawing(s) Number of Sheets _____ ☐ Other (specify) _____

☒ Specification (e.g. description of the invention) Number of Pages 21

Fees Due: Filing Fee of \$220 (\$105 for small entity). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is \$270 (\$130 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT

☐ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check or money order is enclosed to cover the filing fee and application size fee (if applicable). 220.00

☐ Payment by credit card. Form PTO-2038 is attached

☒ The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit

Account Number: 50-1379 A duplicative copy of this form is enclosed for fee processing.

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION COVER SHEET
Page 2 of 2

PTO/SB/16 (10-07)

Approved for use through 06/30/2010. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.	
☒	No.
☐	Yes, the name of the U.S. Government agency and the Government contract number are: _____

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

SIGNATURE <u>/Ronald S. Liu, Reg No 64,170/</u>	Date <u>June 30, 2011</u>
TYPED or PRINTED NAME <u>Ronald S. Liu</u>	REGISTRATION NO. <u>64,170</u> (if appropriate)
TELEPHONE <u>972-583-8512</u>	Docket Number: <u>P34680-US1</u>

Absolute Reference Picture Signaling

1 Background

1.1 Technical Background/Existing Technology

1.2 H.264 / HEVC video compression

H.264 (MPEG-4 AVC) is the state of the art video coding standard. It consists of a block based hybrid video coding scheme that exploits temporal and spatial prediction.

High Efficiency Video Coding (HEVC) is a new video coding standard currently being developed in Joint Collaborative Team - Video Coding (JCT-VC). JCT-VC is a collaborative project between MPEG and ITU-T. Currently, an HEVC Model (HM) is defined that includes large macroblocks (abbreviated LCUs for Largest Coding Units) and a number of other new tools and is considerably more efficient than H.264/AVC.

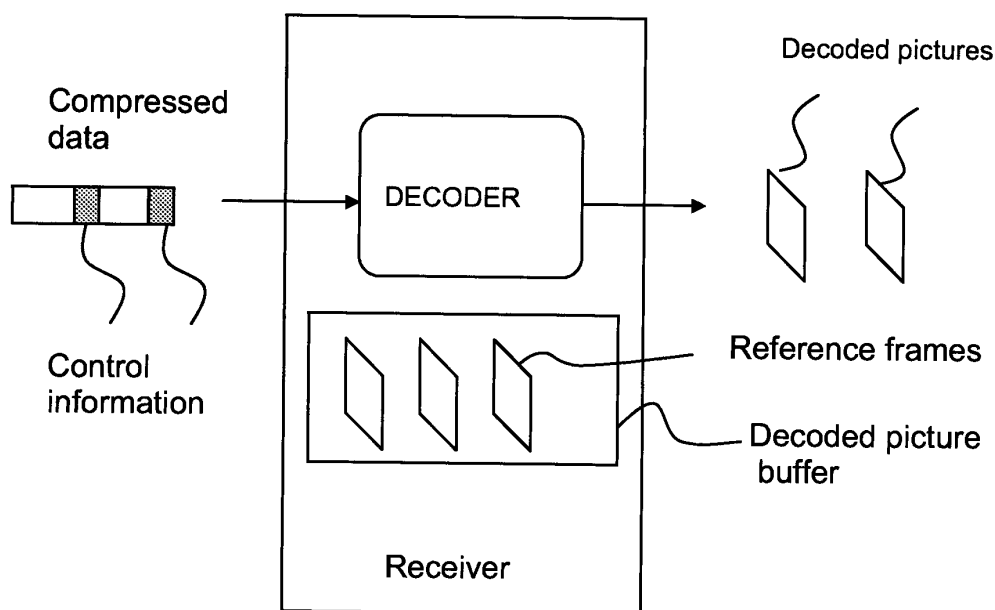


Figure 1a

At a receiver a decoder receives a bit stream representing pictures, i.e. video data packets of compressed data. The compressed data comprises payload and control information. The control information comprises e.g. information of which reference pictures should be stored in a reference picture buffer. This information is a relative reference to previous received pictures. Further, the decoder decodes the received bit stream and displays the decoded picture. In addition, the decoded pictures are stored in reference picture buffer according to the control information. These stored reference pictures are used by the decoder when decoding subsequent pictures.

1.3 Current decoded picture buffer scheme

A working assumption for the processes of decoded picture buffer operations in the working draft of HEVC [1] is that they will be inherited from H.264/AVC to a very large extent. A simplified flow chart of the scheme as it is designed in H.264/AVC is shown in **Figure 1**.

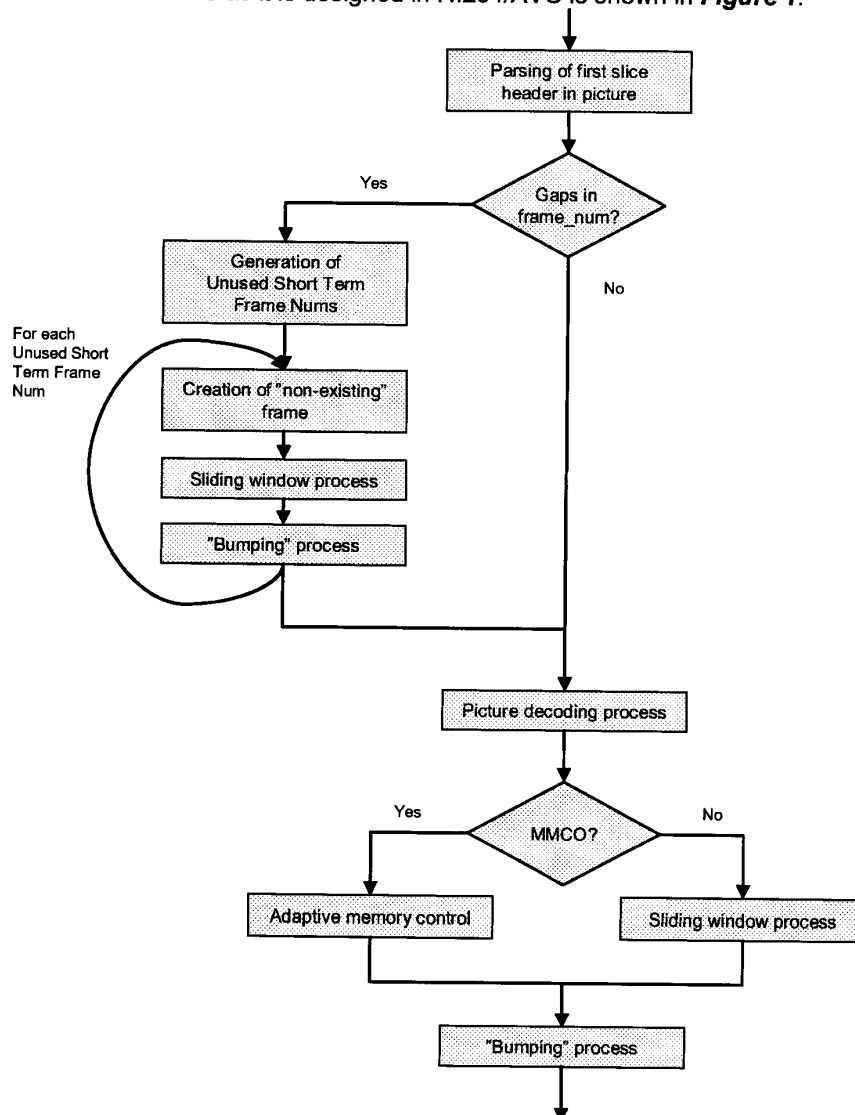


Figure 1 . H.264/AVC reference buffer scheme.

Before the actual decoding of a picture, the frame_num in the slice header is parsed to detect possible gap in frame_num if Sequence Parameter Set (SPS) syntax element

gaps_in_frame_num_value_allowed_flag is 1. The frame_num indicates the decoding order. If a gap in frame_num is detected, "non-existing" frames are created and inserted into the Decoded Picture Buffer (DPB) (as illustrated in figure 1a).

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 contains MMCO commands, adaptive memory control process is applied after decoding of the picture to obtain relative reference to the pictures to be stored in the reference picture buffer; otherwise sliding window process is applied to obtain relative reference to the pictures to be stored in the reference picture buffer. As a final step, the "bumping" process is applied to deliver the pictures in correct order and on time for display.

1.4 Problems with Existing Solutions x

1.4.1 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 constrained intra blocks, intra slices 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.

1.4.2 Sliding window

The first-in-first-out (FIFO) scheme of the sliding window process is not well suited for hierarchical B picture structures. The contribution E399 [2], presents the example of GOP size 16 (15 B pictures in temporal layers 1 – 4). Using sliding window without MMCO requires 8 reference pictures. If using MMCO, or if a more suitable scheme for buffer usage is applied, it is possible to only use 5 reference pictures.

However, the solution proposed in E399 does not guarantee correct operation when gaps in frame_num is allowed. Also it does not minimize the DPB size for all cases, e.g. GOP size 6 with 5B in temporal layers 1-2 as illustrated in **Figure 2**, where 4 reference pictures are required instead of the optimal 3.

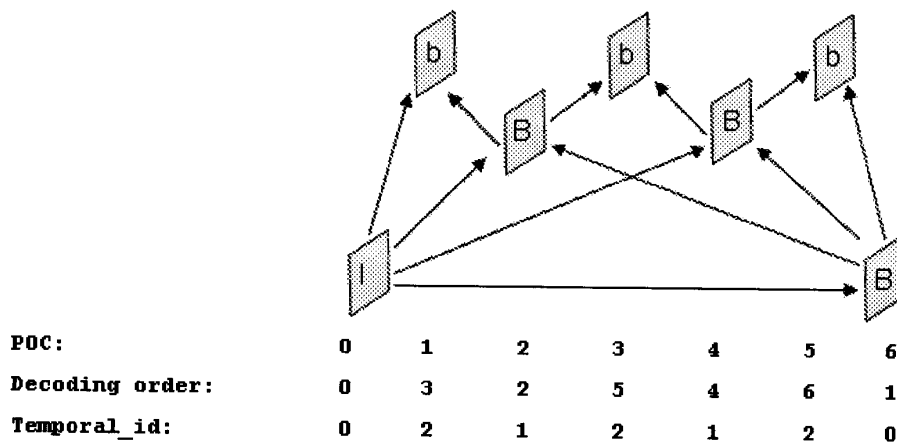


Figure 2. Example coding structure with 5B.

1.4.3 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 MMCOs 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 selection of coding structure and reference picture usage. Consider the example in Figure 3.

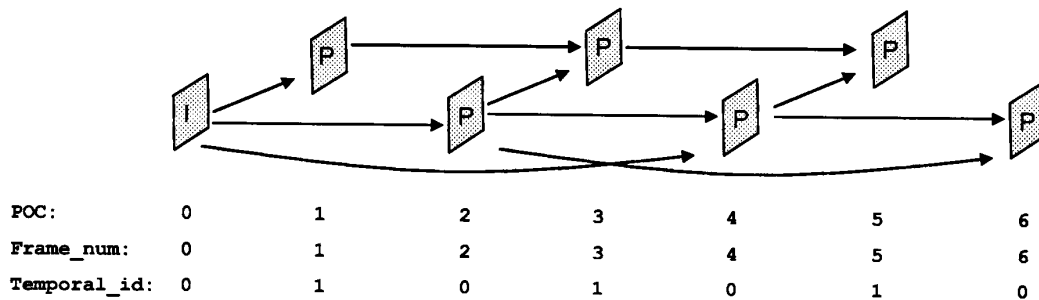


Figure 3. Example coding structure with two temporal layers.

Assuming max_num_ref_frames is 3 even though each picture only uses 2 reference pictures for inter prediction (because each picture must hold one extra picture from the other temporal layer that will be used for inter prediction by the next picture).

In order to have picture 0 and picture 2 available when decoding picture 4, picture 3 must have an explicit reference picture marking 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 it 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 pictures are decoded and when only the lowest layer is decoded, the example coding structure cannot be used for temporal scalability.

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.

1.4.4 Gaps in frame_num and B slices reference picture list initialization

1.4.4.1 Missing POC (picture order count=,display order) value

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)."

1.4.4.2 Missing temporal_id

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.

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 "non-existing" pictures are not assigned temporal_id 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 Summary

A problem with the prior art solutions is that correct decoder reference picture management is dependant on that previously decoded pictures have been correctly received and decoded.

That is solved by specifying which decoded pictures to be used for reference pictures in an absolute (explicit) way instead of a relative (implicit) way.

Another way to put it is that the encoded representation (bitstream) for the current picture contains the information about what pictures to use for reference (reference pictures) independent of the encoded representations of previous pictures.

It can be said that the logical responsibility for maintaining a correct reference picture buffer is moved from the decoder to the bitstream.

The embodiments of the invention are strengthened further by applying buffer operations (e.g. the picture marking process in H.264/AVC where pictures are marked as short-term, long-term and unused-for-prediction) before the pixel decoding of the current picture thus having the explicit buffer description to apply to the current picture. This makes reference picture management even less vulnerable to errors, improves the possibilities for temporal scalability and reduces the overhead introduced by reference picture list modification signaling.

One way to look at it is to say that the information about what reference pictures to use for inter prediction and motion vector prediction for a picture P is included in the control information of picture P.

The embodiments increase the error robustness level of a video decoder. 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.

Using embodiments of the present invention it is possible to apply the optimal buffer scheme (using as few image buffers as possible) for any coding structure with very little overhead. What is needed in the slice header is simply a reference to the correct Buffer Description in the picture parameter set (PPS) or similar. Note that the buffer description can also be put in a picture header or similar, or a parameter set shared between all slices in a picture. The important property is that the buffer description to be used to decode a specific picture is sent with this picture and not with the previous picture in decoding order as in the prior art.

Prior art solutions have problems with temporal scalability for certain coding structures (as described in section 1.3.3) since the information about the reference pictures for the current picture is affect by the removal of a previous picture from the bitstream as the reference picture information is signalled in a relative way. Using embodiments of the present invention, temporal scalability will not be restricted as in prior art since the reference picture for the current picture is

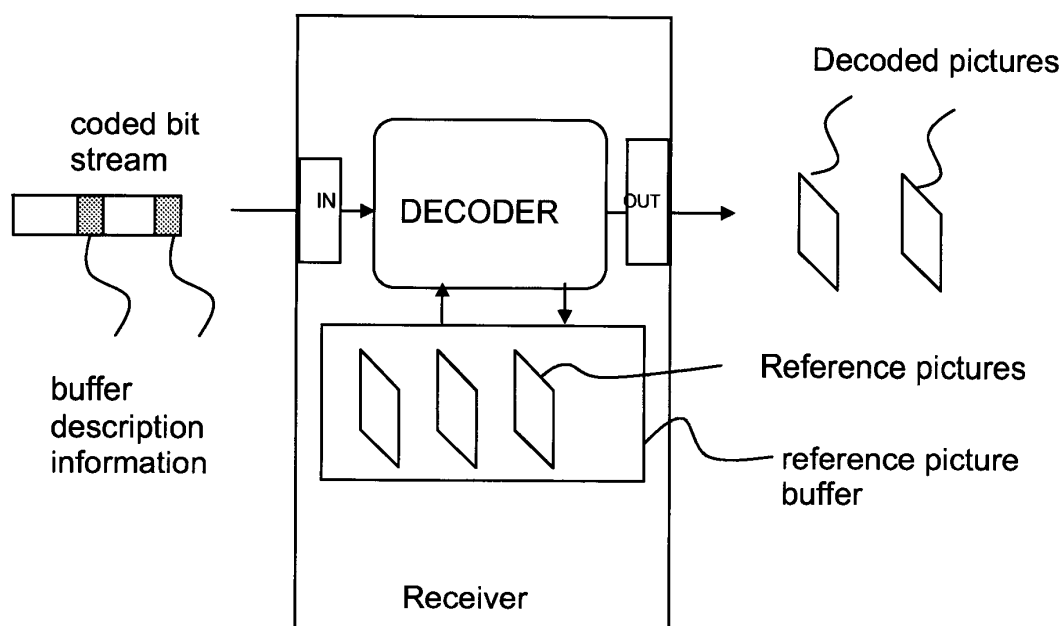
not affected by removal of previous pictures from the bitstream since it is signalled in an absolute way.

According to a first aspect of embodiments of the present invention, a method in a receiver is provided. In the method, a coded data stream is received, buffer description information is detected in the decoded data stream, based on the detected buffer description determine absolute reference to pictures to be used as reference pictures for future decoding, and store the pictures to be used as reference pictures in a reference picture buffer.

According to another aspect: In the method, a coded data stream containing a coded picture is received and buffer description information is detected in the decoded data stream. The buffer description is decoded to determine absolute references to pictures to be used as reference pictures for decoding of said picture, and the buffer description is used to store and release reference pictures in a reference picture buffer]

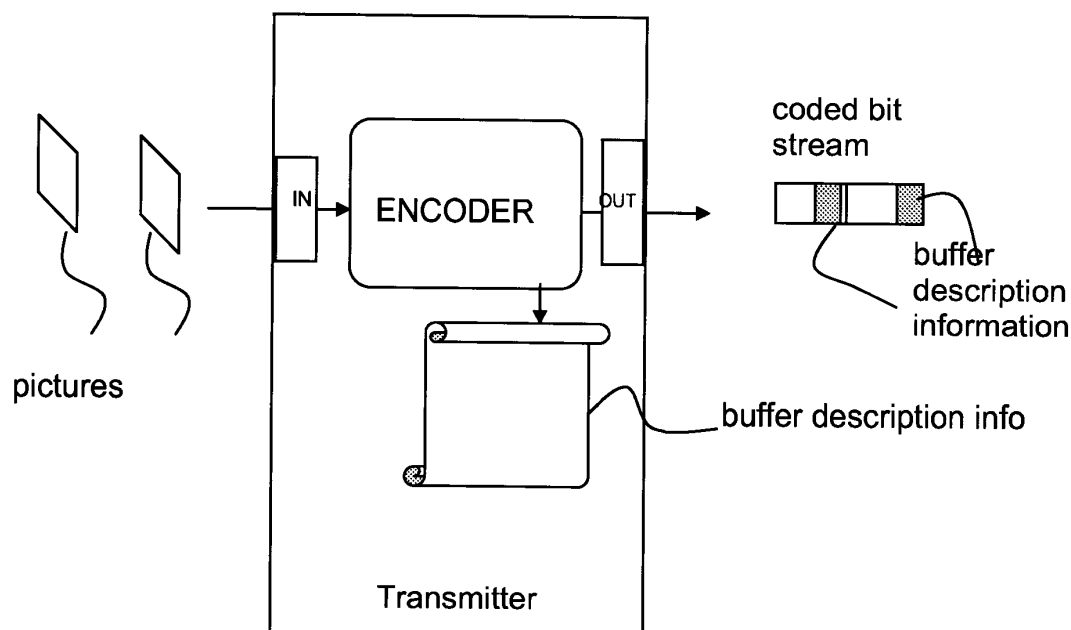
According to a second aspect a method in a transmitter is provided. In the method, an encoder of the transmitter creates buffer description information e.g. a data structure such as a general table with absolute information of which pictures to be used for reference pictures, i.e. stored in a reference picture buffer (also referred to as a decoded picture buffer) to be used for subsequent decoding. At least a part of the buffer description information is inserted in the bit stream encoded by the encoder. I.e. the buffer description information may comprise a table as shown in below (table 2), where a general table is transmitted from the encoder to the decoder and each packet only comprises an entry to be used in the table for determining absolute references to the pictures to be used as reference pictures or where a list is sent in each packet for determining absolute references to the pictures to be used as reference pictures. (If explicit signalling is used, the entire table is not sent in each packet only the information used for that packet i.e. a list corresponding to one row of table 2). The encoder is free to select which pictures to include in the buffer description and may base its selection on any aspect (e.g. the closest pictures in display order). Typically, an encoder selects the buffer descriptions to try to achieve maximum compression efficiency while complying to a set of side constraints. One examples of such a constraint is a maximum number of reference pictures due to memory size. Another example is that the video stream shall be decodable also when certain coded pictures in the bitstream are removed from the bitstream before decoding.

According to a third aspect, a receiver is provided. The receiver comprises an input section for receiving a coded data stream, a decoder for detecting buffer description information and for determining absolute reference to pictures to be used as reference pictures for future decoding based on the detected buffer description, and a buffer for storing the pictures to be used as reference pictures. The receiver further comprises an output section for outputting pictures to be displayed.



The input section is according to an embodiment also configured to receive a data structure such as a table comprising general buffer description information. This data structure can be stored in a memory of the receiver (not shown).

According to a fourth aspect a transmitter is provided. The transmitter comprises an encoder configured to create buffer description information e.g. a data structure such as a general table with absolute information of which pictures to be used for reference pictures, i.e. stored in a reference picture buffer (also referred to as a decoded picture buffer) to be used for subsequent decoding. At least a part of the buffer description information is inserted in the bit stream encoded by the encoder. I.e. the buffer description information may comprise a table as shown in below (table 2), where a general table is transmitted from the encoder to the decoder and each packet only comprises an entry to be used in the table for determining absolute references to the pictures to be used as reference pictures or where a list is sent in each packet for determining absolute references to the pictures to be used as reference pictures. (If explicit signalling is used, the entire table is not sent in each packet only the information used for that packet i.e. a list corresponding to one row of table 2)



3 Detailed description _x

According to an embodiment of the present invention, the indication of the absolute reference to pictures to be used as reference pictures is provided by a Buffer Description. Buffer Description information is provided in control information of the coded pictures provided from the encoder to the decoder. At a minimum, the Buffer Description contains information needed by the decoder to identify the pictures in the reference buffer. Such information is exemplified by a decoding order number (referred to as *frame_num*) or a display order number (referred to as *POC*) or a combination of display order number and an additional id (*POC* and *additional_picture_id*) or any other information that can be used to unambiguously identify the pictures.

In one embodiment, the Buffer Description contains information about each picture *P* that is used by the decoder in the decoding process even if *P* is not used for inter prediction or motion vector prediction (reference frames) or any other prediction (i.e. effectively not used for reference). Such information could include, but is not limited to, decoding order number (*frame_num*), display order number (*POC*), temporal layer information (*temporal_id*), view information (*view_id*).

During the decoding process, it may happen that the Buffer Description contains information about pictures (also referred to as frames) that are not available and therefore can not be used for reference pictures. In one embodiment, a number of "non-existing" pictures is generated by the decoder if the Buffer Description contains information about pictures that are not available in the reference picture buffer. Each such picture *P* can be given values to variables holding information that is used by the decoder in the decoding process even if *P* is not used for inter prediction or motion vector prediction. Such information could include, but is not limited to, decoding order number (*frame_num*), display order number (*POC*), temporal layer information (*temporal_id*), view information (*view_id*).

In addition, pictures may be available in the reference picture buffer but not included in the Buffer Description. Hence, in one embodiment, pictures that are available in the reference picture buffer

but not included in the Buffer Description are removed by the decoder from the reference picture or marked as "unused for prediction".

In one embodiment, the marking of pictures as "unused for prediction" is performed by the decoder prior to the picture decoding process of the picture containing the Buffer Description.

In a further embodiment, zero or more of the pictures that are marked as "unused for prediction" by the decoder according to the Buffer Description are outputted for display by the decoder. One such example process for output is the "bumping process" from H.264/AVC. Output refers to "output for display". What frames to use as reference frames and what frames to output (display) is separated in H.264 and HEVC i.e. a frame can be outputted before it is "removed as reference frame" (by marking it as "unused for reference") or it can be "removed as reference frame" (by marking it as "unused for reference") before it is outputted.

In yet a further embodiment, the marking of zero or more pictures as "unused for prediction" by the decoder is followed by output for display of zero or more pictures marked as "used for prediction" or "unused for prediction". One such example process for output is the "bumping process" from H.264/AVC.

In one embodiment, the generation of "non-existing" pictures is performed by the decoder prior to the picture decoding process of the picture containing the Buffer Description.

In one embodiment, the marking of zero or more pictures is performed by the decoder in a first step. In a second step, zero or more pictures are outputted by the decoder. In a third step zero or more "non-existing" pictures are generated by the decoder.

In an alternative embodiment, the marking of zero or more pictures is performed by the decoder in a first step. Then an iterative process of outputting pictures and generating "non-existing" pictures is performed by the decoder for each picture described in the Buffer Description that is not available in the reference picture buffer.

In one embodiment, the Buffer Description is explicitly signalled to the decoder for the picture it applies to. That explicit signalling might be in the slice header of one, multiple or all slices of one picture. Alternatively the Buffer Description can be signalled in a Slice Parameter Set or a Picture Header or another data structure specified by the standard to which the decoder conforms.

In a further embodiment, this explicit signalling is done by signalling a set of reference picture properties as given above (e.g. frame_num, POC, temporal_id, view_id) relative to the value of these properties as signalled in the current picture.

In one embodiment, one or more Buffer Descriptions are signalled in a data structure so that same Buffer Description can be used for multiple pictures. That signalling can be carried in a Picture Parameter Set, Sequence Parameter Set, a novel parameter set or in another data structure specified by the standard to which the decoder conforms. To specify which Buffer Description to use for the current picture an identifier can be signalled for the current picture. An example of such an identifier is a non negative integer signalled in the slice header(s) of the current picture representing the number of the Buffer Description in the order in which the Buffer Descriptions are signalled. An example of such a data structure is a table (table 2) as shown below.

An example buffer description can be seen in **Table** where 4 slots for picture buffers are used.

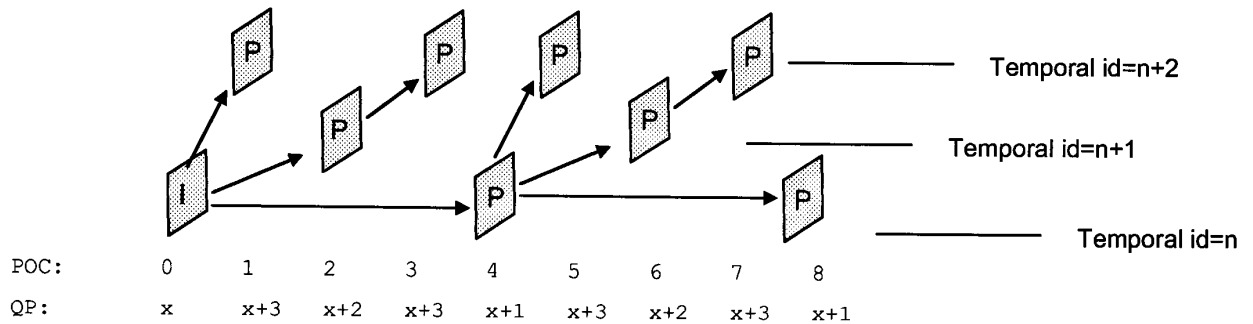


Figure 3 . Coding structure with (sub)GOP size 8.

An IDR picture (An IDR frame is an I frame that removes all dependencies to data transmitted prior to the IDR frame i.e. it marks all reference pictures as "unused for reference") does not need a buffer description (since it empties the buffer) but the other pictures do. Table 2 shows one example of a Buffer Description table where deltaPOC (dP) and temporal_id (tld) are indicated. The table 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.

Entry	Picture buffer 1	Picture buffer 2	Picture buffer 3	Picture buffer 4
0	dP: -1 tld: 2	dP: -2 tld: 1	dP: -4 tld: 0	dP: -8 tld: 0
1	dP: -1 tld: 0	dP: -2 tld: 2	dP: -5 tld: 0	dP: -9 tld: 0
2	dP: -1 tld: 2	dP: -2 tld: 0	dP: -6 tld: 0	dP: -10 tld: 0
3	dP: -1 tld: 1	dP: -2 tld: 2	dP: -3 tld: 0	dP: -7 tld: 0

Table 2. Example Buffer Description. dP for deltaPOC and tld for temporal_id

One example use of this table by the encoder is to for a picture with POC=n, signal table entry (n%4) to the decoder. The reference picture buffer consist of 4 pictures. Which pictures depends on the POC of the current picture and what entry is used. For example if picture with POC 7 uses entry 3 the reference picture buffer will consist of pictures with POC {6, 5, 4, 0}

This table may be created in the encoder and transmitted to the decoder. In the control information of the encoded bit stream the entry number to be used is provided. By using such a table the decoder can obtain absolute information of which pictures to be used as reference pictures by detecting the entry number in the decoded bit stream and using that entry number look up the entry in the table and based on that determine which pictures to be stored in the reference picture buffer.

According to a further embodiment, explicit signalling of Buffer Descriptions and reference signalling to an entry in a general Buffer Description such as an entry in the table above, can be combined by the decoder to form the final Buffer Description for the current picture. One way to combine the explicit signalling and the reference signaling is to join the set of reference pictures described by explicit signalling with the set of reference pictures described by reference signaling to form a joint set of reference pictures. Then one or more reference pictures included in the set of reference pictures described by reference signaling but not included in the explicit signaling are possibly removed from the joint set of reference pictures in order to have a final Buffer Description with no more than max_num_ref_frames number of reference pictures. Preferably pictures are removed in the order that they are listed in the general Buffer Description starting with the last one going forward.

In one embodiment, information available for the current picture is used by the decoder to construct the final Buffer Description for the current from the signalled Buffer Description. Such information includes but is not limited to current POC [POC(curr)] which together with a signalled

deltaPOC can be used to calculate the POC of the reference picture [POC(ref)] as $\text{POC(ref)} = \text{POC(curr)} + \text{deltaPOC}$.

In one embodiment, temporal layer information (temporal_id) is included for each picture in a Buffer Description signalled using $\text{ceil}(\log_2(\text{max_temporal_layers_minus1}))$ bits for signalling of the temporal_id.

In one embodiment, delta display order information (deltaPOC) is included in the Buffer Description encoded with a variable length code for absolute_delta_POC_minus_one and a flag (a single bit) for deltaPOC_sign signalled only if number_of_reorder_frames > 0, otherwise the sign is inferred to be negative.

3.1.1 Extension

In one embodiment, the Buffer Description may contain information that is used by the decoder in reference picture list initialisation or reference picture list modification or reference picture list combination. One example is that the order in which the pictures are listed in a Buffer Description can be used as the initial order for one of the reference picture lists in reference picture list initialisation.

In one embodiment, restrictions are specified for the pictures included in the Buffer Descriptions. An example restriction is that; a picture A with temporal_id tld(A) described in the Buffer Description for a picture B with temporal_id tld(B) may not be a "non-existing" picture if $\text{tld(A)} < \text{tld(B)}$.

In one embodiment, it is possible to signal changes to the properties of pictures in the reference picture buffer through Buffer Descriptions. One example is to change temporal layer and/or display order of pictures through signalling the new value for temporal layer and/or display order in the Buffer Description.

3.2 Description of the proposal to HEVC standardization

These are the changes that we propose:

- Move decoder picture buffer marking from after the decoding of a picture to after parsing of the first slice header of a picture
- Add explicit signaling of the reference pictures (Buffer Description) in each slice header, containing POC and temporal_id of all reference pictures available for prediction in the DPB
- Include Buffer Descriptions in the Picture Parameter Set to minimize the overhead of signaling available reference pictures (Buffer Descriptions) in slice header
- Remove all MMCO commands
- Remove the sliding window operation
- Remove syntax elements frame_num and gaps_in_frame_num
- 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)
- Remove the conventional concept of long term pictures

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

We propose a scheme where all decoded picture buffer operations are applied after parsing of the first slice header of a picture but before picture decoding, using a description of the decoded picture buffer as illustrated in **Figure 4**. 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 large conceptual changes to the decoding process. In H.264/AVC and current design of HEVC, relative operations are given to the decoder either

implicitly (sliding window) or explicitly (MMCO) and the decoder is responsible of applying these relative 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 signalled relative 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).

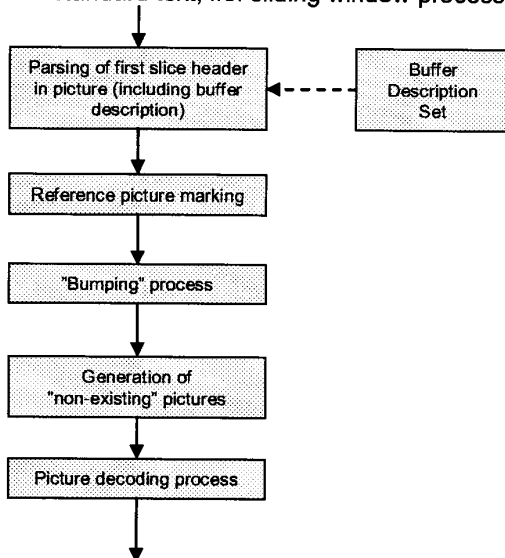


Figure 4 . Proposed scheme.

3.2.1 Buffer description

According to an embodiment, the buffer description contains the *delta_POC*, *temporal_id* and *additional_picture_id* of all reference pictures (in the DPB) in order to provide an absolute reference to the pictures to be used as reference pictures. The *delta_POC* is used to calculate the POC of a reference picture as $POC(ref) = POC(current) + delta_POC$. Pictures will be identified by the pair; POC and *additional_picture_id*. *Temporal_id* is included in the description to enable correct reference picture list modification in the case of lost or removed pictures (e.g. temporal scalability).

All pictures in the DPB that are not a part of the buffer description shall be marked as "unused for prediction". The idea is not restricted to the codewords *delta_POC*, *temporal_id* and *additional_picture_id*. Any codeword that is associated with a picture and used in the reference picture handling is included in the buffer description, either relative to the value of the current picture (e.g. POC and *delta_POC*) or absolute (e.g. *temporal_id*).

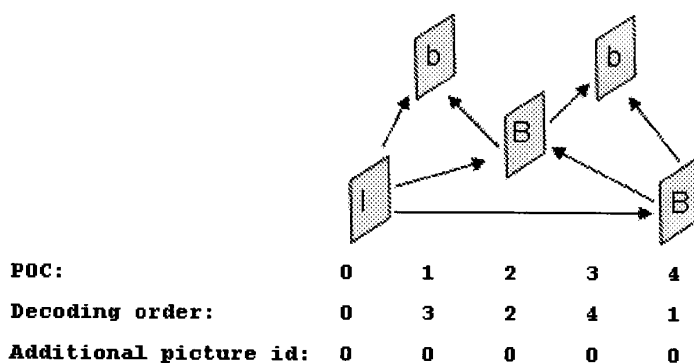


Figure 5 . Example coding structure.

In the example in **Figure 5**, 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 it will be made "unused for prediction".

In the example in **Figure 5**, additional_picture_id is 0 for all pictures. It is recommended to set additional_picture_id to 0 unless there are two different pictures available for reference with the same POC (due to POC wrap-around). If POC is the same for two different pictures the additional_picture_id shall be different in order to not (unconditionally) replace the older picture with the newer one. Additional_picture_id is introduced in order to provide all the functionalities and possibilities that are provided by long term pictures in H.264/AVC.

Consider an example where an encoder would like to use a picture A with POC 14, and additional_picture_id set to 0, as a "long term picture". The encoder must then make sure that the picture A is included in the buffer description in the slice header of all pictures following A for as long as the encoder wants picture A to be available for reference. If the encoder would like to encode another picture with POC 14 (due to POC wrap-around) while A is still available for reference, the encoder should select another value of additional_picture_id, for example 1, to avoid that A is replaced as a reference picture.

Pictures that are signaled in the buffer description but does not exist in the DPB shall be marked as "non-existing". However since the POC and the temporal_id of such pictures are included in the buffer description, no explicit signaling of reference picture list modification is required. If the "non-existing" picture is not used for prediction, the decoding process will proceed correctly without extra signaling (and without the overhead incorporated with signaling reference picture list modifications explicitly).

3.2.1.1 Buffer description entries in Picture Parameter Set

To reduce the bit overhead of signaling the buffer descriptions explicitly in the slice header it is proposed to introduce buffer description entries in the Picture Parameter Set. These descriptions can be used for multiple slices/pictures in the same sequence and thus reduce the number of bits required per picture.

For each buffer description entry in the Picture Parameter Set the delta_POC and temporal_id of all reference pictures should be present. The additional_picture_id is inferred to be 0 for pictures described by an entry in the PPS.

3.2.1.2 Removal of frame_num and gaps_in_frame_num_allowed

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 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 combination of POC and additional_picture_id 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.

3.2.2 Implications on the decoding process

An IDR picture does not need to send a buffer description since it does not use any previous pictures for reference and removes all reference pictures from the DPB.

For all other types of pictures (not IDR) a buffer description is **required** in order to construct the correct decoded picture buffer before decoding of the current picture. Pictures are identified by the pair POC (calculated as $POC(current) + \delta_{poc}$) and additional_picture_id. For pictures that will be used for reference for a short period of time (short term reference pictures) it is recommended to set additional_picture_id to 0, in order to minimize the bit overhead. For pictures that will be used for reference for a long period of time (long term pictures) it is recommended that the encoder makes sure that no two pictures with the same POC and the same additional_picture_id are available for prediction at the same time.

The total number of reference pictures is indicated by the syntax element max_num_ref_frames in the SPS.

A buffer description entry in a Picture Parameter Set may have at most max_num_ref_frames pictures in its description. It may contain fewer pictures in which case the remaining ones are interpreted as "Empty".

A picture that does not contain a reference to a buffer description must contain explicit description of all reference pictures (otherwise the reference pictures will be marked as "unused for reference").

In order to reduce bit overhead we propose that it should be possible to both have a reference to a Buffer Description in a Picture Parameter Set and to explicitly signal Buffer Descriptions in the same slice header. If a slice header contains both a reference and explicit description(s), the descriptions (reference pictures) fetched from the PPS will be replaced by the explicit descriptions starting from the end of the list in the order in which they were described in the PPS.

A picture that is used for reference for a long period of time (long term reference picture) is preferably indicated by explicit description 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 Descriptions in the PPS.

3.2.3 Implications on temporal scalability

The proposed scheme implies that temporal down-switching is always possible. Each temporal layer (together with the lower layers) will constitute a sub-sequence. Such sub-sequences need no explicit signaling.

3.2.4 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 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 6**. 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.

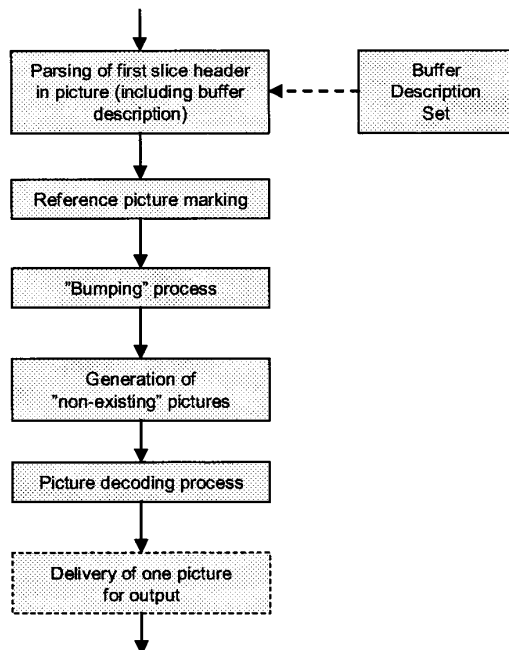


Figure 6 . Alternative scheme.

4 Proposed syntax

Added syntax elements are highlighted (in yellow) and removed syntax elements are marked with strikethrough.

slice_header() {	Descriptor
first_tb_in_slice	ue(v)
entropy_slice_flag	u(1)
if(!entropy_slice_flag) {	
slice_type	ue(v)
pic_parameter_set_id	ue(v)
~~frame_num~~	~~u(v)~~
additional_picture_id	ue(v)
if(IdrPicFlag)	
idr_pic_id	ue(v)
if(pic_order_cnt_type == 0)	
pic_order_cnt_lsb /*	u(v)
if(!IdrPicFlag) {	
buffer_description_reference_flag	u(1)
if(buffer_description_reference_flag == 1) {	
buffer_description_id	ue(v)
}	
if(buffer_description_reference_flag == 0) {	
number_of_explicitly_signaled_pictures	ue(v)
for(number_of_explicitly_signaled_pictures) {	
additional_picture_id	ue(v)
if(num_reorder_frames > 0)	
delta_POC_sign	u(1)
absolute_delta_POC_minus_one	ue(v)
temporal_id	u(v)
}	
}	
}	
}	
if(slice_type == P slice_type == B) {	
num_ref_idx_active_override_flag	u(1)
if(num_ref_idx_active_override_flag) {	
num_ref_idx_l0_active_minus1	ue(v)
if(slice_type == B)	
num_ref_idx_l1_active_minus1	ue(v)
}	
}	
ref_pic_list_modification()	
ref_pic_list_combination()	
~~if(nal_ref_idc != 0)~~	
~~do_ref_pic_marking()~~	
if(entropy_coding_mode_flag && slice_type != I)	
cabac_init_idc	ue(v)
slice_qp_delta	se(v)
if(adaptive_loop_filter_enabled_flag)	
alf_param()	

if(deblocking_filter_control_present_flag) {	
disable_deblocking_filter_idc	
if(disable_deblocking_filter_idc != 1) {	
slice_alpha_c0_offset_div2	
slice_beta_offset_div2	
}	
}	
if(slice_type == B)	
collocated_from_l0_flag	u(1)
} else	
if(entropy_coding_mode_flag && slice_type != I)	
cabac_init_idc	ue(v)
}	

4.1.1 Picture Parameter Set

	Descriptor
pic_parameter_set_rbsp() {	
pic_parameter_set_id	ue(v)
seq_parameter_set_id	ue(v)
entropy_coding_mode_flag	u(1)
num_ref_idx_l0_default_active_minus1	ue(v)
num_ref_idx_l1_default_active_minus1	ue(v)
pic_init_qp_minus26 /* relative to 26 */	se(v)
constrained_intra_pred_flag	u(1)
number_of_buffer_descriptions	ue(v)
for(number_of_buffer_descriptions){	
number_of_pictures_in_buffer_description	ue(v)
for(number_of_pictures_in_buffer_description) {	
if(num_reorder_frames > 0)	
delta_POC_sign	u(1)
absolute_delta_POC_minus_one	ue(v)
temporal_id	u(v)
}	
}	
rbsp_trailing_bits()	
}	

5 Proposed semantics

8.2 Slice decoding process

8.2.2 Decoding process for buffer description application

This process is invoked at the beginning of the decoding process for each I, P, or B picture, prior to decoding of the first Coding Unit in the first slice of the picture and prior to the decoding process for reference picture list construction as specified in subclause 8.2.3.

The process may result in marking one or more picture as "unused for reference" as described in 8.2.2.1.

The process may result in the construction of "non-existing" pictures as described in 8.2.2.2.

A reference picture is identified for use in the decoding process by the variables POC and additional_picture_id.

A reference picture r in the reference picture buffer is said to be a part of a buffer description if and only if $POC(r) = POC(curr) + \delta_{POC}(i)$ and $additional_picture_id(r) = additional_picture_id(i)$ for any i in the range of 0 to number_pictures_in_buffer_description.

8.2.2.1 Sequence of operation for buffer description application

Buffer description application proceeds in the following ordered steps:

1. Depending on whether the current picture is an IDR picture, the following applies.
 - If the current picture is an IDR picture, all reference pictures are marked as unused for reference.
 - Otherwise (the current picture is not an IDR picture), the following ordered steps apply:
 - I. The buffer description information in the slice header of the first slice of the current picture is used to create a list of reference pictures, referred to as the buffer description as described in 8.2.2.2.
 - II. All reference pictures (in the reference picture buffer) that are not a part of the buffer description are marked as "unused for reference".
 - III. If one or more pictures that are included in the buffer description are not available for reference (i.e. existing in the reference buffer marked as "used for reference") the process described in 8.2.2.2 is invoked.
2. The slices of the current picture are decoded.

8.2.2.2 Creation of a buffer description

If buffer_description_reference_flag is 1, a buffer description is created setting deltaPOC to the deltaPOC of the PPS buffer description entry identified by buffer_description_id, temporal_id to the temporal_id of the PPS buffer description entry identified by buffer_description_id and additional_picture_id to 0 for each picture i in the range from 0 to number_of_pictures_in_buffer_description-1.

If buffer_description_reference_flag is 0, a buffer description is created setting deltaPOC to the deltaPOC of the explicitly signaled buffer description, temporal_id to the temporal_id of the PPS explicitly signaled buffer description and additional_picture_id to additional_picture_id of the explicitly signaled buffer description for each picture i in the range from 0 to number_of_explicitly_signaled_pictures -1.

8.2.2.3 Creation of non-existing pictures

This process is invoked for each picture that is included in the buffer description of the current picture but is not available in the decoded picture buffer. That is if the buffer description of the current picture contains an entry i , with a $POC(i)$ and $additional_picture_id(i)$ such that there is no picture j in the reference picture buffer marked as "used for reference" with $POC(j) == POC(i)$ and $additional_picture_id(j) == additional_picture_id(i)$.

For each such description a "non-existing" picture is generated with POC set to $POC(i)$, additional_picture_id set to $additional_picture_id(i)$, temporal_id set to $temporal_id(i)$ and marked as "non-existing" and "used for reference". The sample values of the generated frames may be set to any value. These generated frames which are marked as "non-existing" shall not be referred to in the inter prediction process.

The encoder comprises at least an input section (IN), an output section (OUT), a processor and one or more memory sections.

The encoder receives pictures to be encoded via the input section and the processor is configured to perform the above described procedures. The functionality of the processor may be realized by software code portions stored in a memory section. A memory section may also store buffer description information. When the pictures are coded, the coded pictures including respective index values are transmitted to the decoder via the output section.

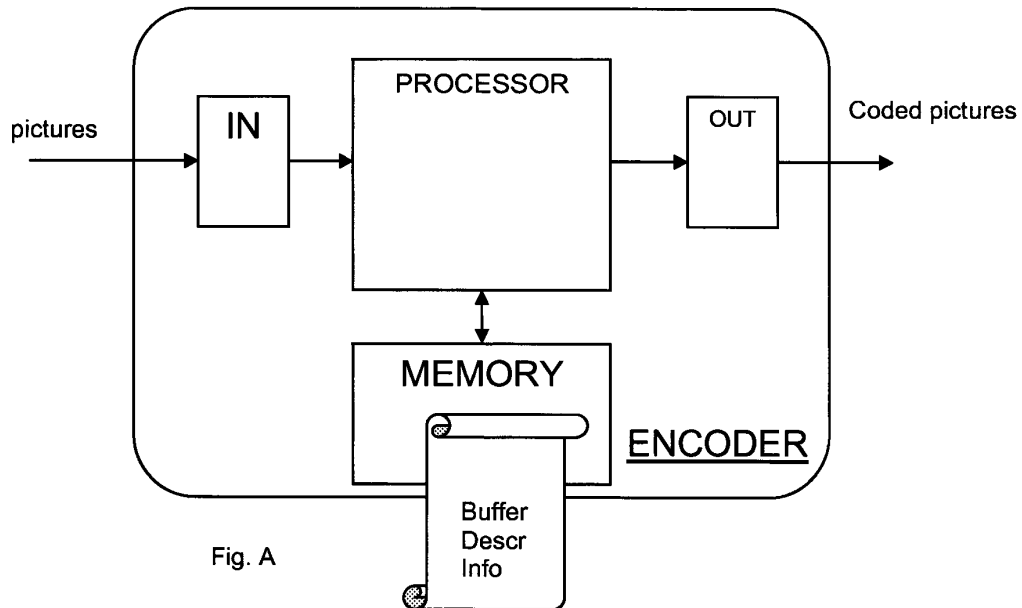
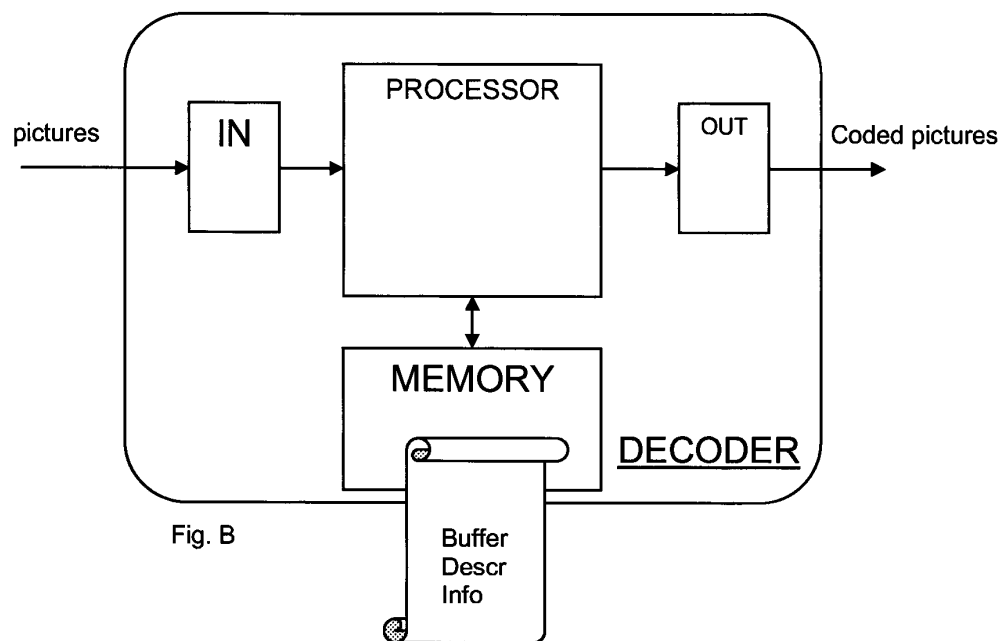


Fig. A

Accordingly, the above described procedure is performed by a decoder. An example of a decoder is shown in figure B.

The decoder comprises at least an input section (IN), an output section (OUT), a processor and one or more memory sections.

The decoder receives encoded pictures via the input section and the processor is configured to perform the above described procedures. The functionality of the processor may be realized by software code portions stored in a memory section. A memory section may also store buffer description information. When the pictures are decoded, the decoded pictures are sent for display.



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.
http://phenix.int-evry.fr/jct/doc_end_user/documents/5_Geneva/wg11/JCTVC-E603-v7.zip
- [2] JCTVC-E399, Sliding Window for Temporal Scalability, Y. Chen, M. Karczewicz.
http://phenix.int-evry.fr/jct/doc_end_user/documents/5_Geneva/wg11/JCTVC-E399-v1.zip