



- (51) **International Patent Classification:**
H04N 21/2343 (2011.01) *H04N 21/845* (2011.01)
H04N 21/4402 (2011.01) *H04N 7/26* (2006.01)
H04N 21/647 (2011.01)
- (21) **International Application Number:**
PCT/SE2012/051327
- (22) **International Filing Date:**
29 November 2012 (29.11.2012)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/588,800 20 January 2012 (20.01.2012) US
- (71) **Applicant:** TELEFONAKTIEBOLAGET L M ERICSSON (PUBL) [SE/SE]; SE-164 83 Stockholm (SE).
- (72) **Inventors:** SAMUELSSON, Jonatan; Stadshagsplan 1, S-11250 Stockholm (SE). Rickard, Sjöberg; Robert Almströmgatan 9A, S-11336 Stockholm (SE).
- (74) **Agent:** EGRELIUS, Fredrik; Ericsson AB, Patent Unit Kista DSM, SE-16480 Stockholm (SE).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on next page]

(54) **Title:** SUB-BITSTREAM EXTRACTION

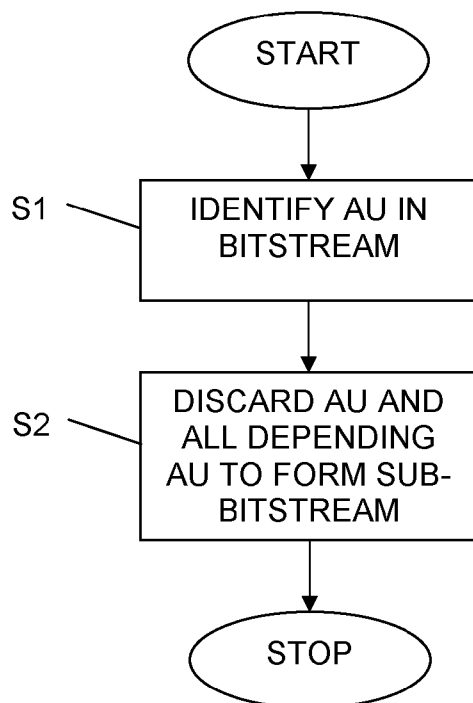


Fig. 4

(57) **Abstract:** A sub-bitstream extraction process is applicable to a bitstream of multiple access units comprising a respective encoded picture (2) of a video sequence (1). An access unit among the bitstream is identified and discarded. In addition, all access units in the bitstream that can depend on this access unit are discarded to form a sub-bitstream. The sub-bitstream will not have any access unit that can depend on the identified access units and will be a specification-conforming bitstream.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG). **Published:**

— *with international search report (Art. 21(3))*

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

SUB-BITSTREAM EXTRACTION

TECHNICAL FIELD

The embodiments generally relate to processing of bitstreams, and in particular to sub-bitstream
5 extraction in connection with layered streams of pictures of a video sequence.

BACKGROUND

H.264 (Moving Picture Experts Group-4 (MPEG-4) Advanced Video Coding (AVC)) is the state of the
art video coding standard. It consists of a block based hybrid video coding scheme that exploits
10 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 International Telecommunication Union Telecommunication
standardization sector (ITU-T). Currently, an HEVC Working Draft (WD) is defined that includes a
number of new tools and is considerably more efficient than H.264/AVC. HEVC also defines a
15 *temporal_id* for each picture, corresponding to the temporal layer the picture belongs to. *temporal_id* is
also present in Scalable Video Coding (SVC), the scalability extension of H.264/AVC.

An HEVC bitstream consists of Network Abstraction Layer (NAL) units which are grouped together in
access units. Each access unit contains a picture associated with a decoding order value and an
20 output order value.

It is said that a bitstream conforms to a standard if it fulfills the requirements for bitstream conformance.
For HEVC the bitstream conformance requirements can be summarized as:

25 "It is a requirement of bitstream conformance that the bitstream shall be constructed according to the
syntax, semantics, and constraints specified in this Specification outside of this annex. It is a
requirement of bitstream conformance that the first coded picture in a bitstream shall be a Random
Access Point (RAP) picture, i.e. an Instantaneous Decoder Refresh (IDR) picture or a Clean Random
Access (CRA) picture or a Broken Link Access (BLA) picture." "For conforming bitstreams, all of the
30 following conditions shall be fulfilled." HEVC lists ten conditions that must be fulfilled by a conforming
bitstream. These ten conditions are listed in the attached Annex A.

Correspondingly, for H.264/AVC the bitstream conformance requirement can be summarized as:

1. "The bitstream is constructed according to the syntax, semantics, and constraints specified in this Recommendation | International standard."

2. "For conforming bitstreams, all of the following conditions shall be fulfilled..." H.264/AVC lists seven
5 such conditions which are found in Annex B.

The temporal layers are ordered and have the property that a lower temporal layer never depends on a higher temporal layer. Thus, higher temporal layers can be removed without affecting the lower temporal layers. The removal of temporal layers can be referred to as temporal scaling.

10

In SVC a sub-bitstream extraction process is defined specifying that a conforming bitstream from which all NAL units with *temporal_id* higher than a defined value are removed shall also be a conformant bitstream.

15 The reason to extract a sub-bitstream can, for example, be to adapt the bitrate of a relayed bitstream in a network node based on changes in network conditions.

The existing sub-bitstream extraction processes, e.g. in SVC, have two major problems:

20 The sub-bitstream extraction process only covers the case when an entire sequence is scaled to a specific layer.

The sub-bitstream extraction process does not give any guarantees to an encoder or a network node regarding scalability. On the contrary, it puts requirements on the bitstream that the encoder and
25 network node must fulfill in order to produce a conforming bitstream. Thus, the encoder must check each sub-bitstream for bitstream conformance. A network node that extracts a sub-bitstream must check the resulting bitstream for bitstream conformance. This makes flexible encoders and network nodes very complex.

30

SUMMARY

It is a general objective to provide an efficient processing of bitstreams.

It is a particular embodiment to enable sub-bitstream extraction in connection with layered streams or sequences of pictures.

These and other objectives are met by embodiments disclosed herein.

An aspect of the embodiments relates to a sub-bitstream extraction method applicable to a bitstream of multiple access units comprising a respective encoded picture of a video sequence. The method comprises identifying an access unit among the bitstream comprising the multiple access units. This access unit is discarded and also all access units in the bitstream that can depend on this access unit. As a consequence a sub-bitstream having no access unit that can depend on the identified and discarded access unit is formed.

10

A related aspect of the embodiments defines a sub-bitstream extracting device. The sub-bitstream extracting device comprises an access unit identifier configured to identify an access unit among a bitstream of multiple access units comprising a respective encoded picture of a video sequence. An access unit discarder is configured to discard the access unit and all access units in the bitstream that can depend on the access unit to form a sub-bitstream having no access unit that can depend on the access unit identified by the access unit identifier.

20

Further relates aspect of the embodiments define a decoder and an encoder comprising the sub-bitstream extracting device and a mobile device comprising such a decoder and/or encoder.

Another aspect of the embodiments relates to sub-bitstream extracting method in a network node. The method comprises discarding access units of a bitstream of multiple access units comprising a respective encoded picture of a video sequence. Each picture of the video sequence has a respective layer identifier identifying a respective layer. The discarding of access units comprises discarding access units with a layer identifier higher than a selected layer identifier X from the bitstream to form a sub-bitstream. The method also comprises relaying, to form an updated bitstream, access units with layer identifier higher than X starting from a layer switching point with a layer identifier equal to X or X+1 in the bitstream. The discarding of access units and the relaying of access units are performed without checking the sub-bitstream and the updated bitstream for bitstream conformance.

30

Another related aspect of the embodiments defines a network node comprising an access unit discarder configured to discard access units of a bitstream of multiple access units comprising a respective encoded picture of a video sequence. Each picture of the video sequence has a respective layer identifier identifying a respective layer. The access unit discarder is configured to discard access

units with layer identifier higher than a selected layer identifier X from the bitstream to form a sub-bitstream. The network node also comprises a relaying unit configured to relay, to form an updated bitstream, access units with layer identifier higher than X starting from layer switching point with a layer identifier equal to X or X+1 in the bitstream. The access unit discarder is configured to discard the
5 access units and the relaying unit is configured to relay the access units without checking the sub-bitstream and the updated bitstream for bitstream conformance.

The embodiments enable a flexible sub-bitstream extraction and simplify the sub-bitstream extraction by automatically outputting sub-bitstreams that are conforming bitstreams without the need for any
10 computationally complex verification for bitstream conformance.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention, together with further objects and advantages thereof, may best be understood by making reference to the following description taken together with the accompanying drawings, in which:

15

Fig. 1 is a schematic illustration of a video sequence of pictures comprising one or more slices;

Fig. 2 is an illustration of a data packet comprising a NAL unit;

20 Fig. 3 is an illustration of an encoded representation of a slice;

Fig. 4 is a flow diagram of a sub-bitstream extraction method according to an embodiment;

Fig. 5 is a flow diagram illustrating an additional, optional step of the method in Fig. 4 according to an
25 embodiment;

Fig. 6 is a schematic block diagram of a sub-bitstream extracting device according to an embodiment;

Fig. 7 is a schematic block diagram of a decoder according to an embodiment;

30

Fig. 8 is a schematic block diagram of a decoder according to another embodiment;

Fig. 9 is a schematic block diagram of a receiver according to an embodiment;

Fig. 10 is a schematic block diagram of an encoder according to an embodiment;

Fig. 11 is a schematic block diagram of an encoder according to another embodiment;

5 Fig. 12 is a schematic block diagram of a transmitter according to an embodiment;

Fig. 13 is a flow diagram of a bitstream extraction method according to another embodiment; and

Fig. 14 is a schematic block diagram of a network node according to an embodiment.

10

DETAILED DESCRIPTION

Throughout the drawings, the same reference numbers are used for similar or corresponding elements.

The embodiments generally relate to processing of bitstreams, and in particular to sub-bitstream
15 extraction in connection with layered streams of pictures. The embodiments therefore enable forming
sub-bitstreams from a bitstream of coded pictures of a video sequence or stream. Sub-bitstreams are
formed by removing selected portions of the bitstream in a particular way. The embodiments perform
such sub-bitstream extraction in a computationally efficient manner, thereby significantly simplifying
forming sub-bitstreams as compared to prior art techniques requiring verifying bitstream conformance
20 for each formed sub-bitstream.

Hence, the prior art sub-bitstream extraction process requires checking a resulting sub-bitstream for
bitstream conformance. Bitstream conformance is well-known in the art and implies that the bitstream
or sub-bitstream is constructed according the syntax, semantics and constraints specified in the
25 relevant video coding standard or specification, such as the H.264/AVC specification or the HEVC
specification. In addition, for HEVC conforming bitstreams the ten conditions specified in Annex A
should be fulfilled, whereas a H.264/AVC conforming bitstream should fulfill the seven conditions
specified in Annex B.

30 This means that the prior art sub-bitstream extraction processes need to check that the formed sub-
bitstream fulfills the relevant conditions and is constructed according to the syntax, semantics and
constraints specified in the relevant video coding specification. This makes the extraction process
computationally complex.

The present embodiments, in clear contrast, relate to a sub-bitstream extraction process that forms a sub-bitstream that will be a conforming bitstream, thereby relaxing the need for any checking or verification of bitstream conformance. This is a major advantage as compared to the prior art.

- 5 A further advantage of the sub-bitstream extraction process is that it is not limited to only cover the case where the entire sequence is scaled to a specific layer, such as a temporal layer. In clear contrast, it is possible to remove parts of a sequence without having to remove all parts with the same or higher layer identifier, such as *temporal_id*, as the removed parts.
- 10 In a basic concept the sub-bitstream extraction is done through a process that given the input to remove a specific picture from a bitstream all pictures that might be affected by that picture are also removed.

Various embodiments will now be further described herein with reference to the drawings.

- 15 Fig. 1 is a schematic overview of a portion of a video sequence or stream 1 of pictures 2. A picture 2 in HEVC and H.264/AVC is portioned into one or more slices 3, where each slice 3 is an independently decodable segment of a picture 2. This means that if a slice 3 is missing, for instance got lost during transmission, the other slices 3 of that picture 2 can still be decoded correctly. Generally, in order to make slices 3 independent, they should not depend on each other. Hence, in a particular embodiment,
- 20 no bitstream element of a slice 3 is required for decoding any element of another slice 3.

In HEVC and H.264/AVC a coded video sequence 1, i.e. bitstream, comprises Network Abstraction Layer (NAL) units 11 as illustrated in Fig. 2. Basically, one NAL unit 11 comprises either a slice with a corresponding slice header including control information for that slice or the NAL unit 11 comprises a

25 parameter set. The parameter set comprises control information.

A NAL unit 11 as output from an encoder is typically complemented with headers 12 to form a data packet 10 that can be transmitted as a part of a bitstream from the encoder to the decoder. For instance, Real-time Transport Protocol (RTP), User Datagram Protocol (UDP) and Internet Protocol

30 (IP) headers 12 could be added to the NAL unit 11. This form of packetization of NAL units 11 merely constitutes an example in connection with video transport. Other approaches of handling NAL units 11, such as file format, MPEG-2 transport streams, MPEG-2 program streams, etc. are possible.

Examples of parameter sets that could be carried in NAL units 11 include Adaptation Parameter Set (APS), Picture Parameter Set (PPS) and Sequence Parameter Set (SPS). APS comprises control information valid for more than one slice. The control information may differ between the slices. PPS comprises control information valid for several pictures, and may be the same for multiple pictures of the same video sequence. SPS comprises control information valid for an entire video sequence. Also Video Parameter Set (VPS) is available in HEVC.

As shown in Fig. 3 an encoded representation 20 of a slice 3 comprises a slice header 21, which independently provides all required data for the slice 3 to be independently decodable. An example of a data element present in the slice header 21 is the slice address, which is used by the decoder to know the spatial location of the slice. Another example is the slice quantization delta which is used by the decoder to know what quantization parameter to use for the start of the slice. The encoded representation 20 also comprises, in addition to the slice header 21, slice data 22 that comprises the encoded data of the particular slice 3, e.g. encoded color values of the pixels in the slice 3.

15

A bitstream consists of NAL units 11, which are grouped together in so-called access units. Each access unit contains a picture 2. Hence, an access unit is a set of NAL units 11 that are associated with each other, are consecutive in a decoding order of the video sequence 1 and contain exactly one coded picture. An access unit may, in addition to the NAL units 11 with coded slice data 22 of the coded picture, also contain other NAL units not containing such encoded representations of slices, such as the above mentioned parameter sets.

Fig. 4 is a flow diagram illustrating a sub-bitstream extraction method, also denoted bitstream extraction method, which is applicable to a bitstream of multiple access units comprising a respective encoded picture of a video sequence. The method generally starts in step S1. Step S1 comprises identifying an access unit (AU) in or among the bitstream comprising multiple access units. This identified access unit is discarded in step S2 together with all access units in the bitstream that can depend on the access unit identified in step S1. The result of step S2 is the formation of a sub-bitstream having no access unit that can depend on the access unit in step S1.

30

In an embodiment, a notification or list comprising an identifier of the access unit or indeed of the coded picture carried by the access unit is typically used in step S1 to identify the relevant access unit in the bitstream. Hence, the identification in step S1 is typically performed based on an access unit identifier

or a picture identifier, such as in the form of a Picture Order Count (POC) identifier, associated with and enabling identifying of the relevant access unit in the bitstream.

The optional identifier used in step S1 could be received from another unit or device than the particular unit or device that performs the sub-bitstream extraction method. Alternatively, the identifier could be provided or determined by the unit or device that performs the sub-bitstream extraction method, for instance based on feedback information with regard to available bandwidth or bitrate requirements for the current video sequence. Thus, the sub-bitstream extraction could be performed based on changing or varying bandwidth or bitrate levels, or indeed based on feedback information from the decoder or the device comprising the decoder. In the latter case, the user of the device could have selected to use a lower quality level for the display of the video, thereby merely requiring the decoding of a sub-bitstream. In such a case, the device or decoder could provide information indicating that a sub-bitstream should be formed from the bitstream.

In a particular embodiment step S2 involves identifying all access units in the bitstream that can depend on the particular access unit identified in step S1. These identified access unit(s) is(are) then discarded together with the particular access unit in step S2.

An access unit in the bitstream that can depend on the access unit identified in step S1 is identified as directly or indirectly depending on or referencing the access unit identified in step S1.

Video coding, such as in HEVC and H.264/AVC, involves two main coding and decoding modes generally referred to as inter coding that uses inter prediction and intra coding that uses intra prediction. In the former case, i.e. inter coding, a picture or a slice of a picture is encoded and decoded with reference to and based on at least one other picture in the bitstream. In the case of a predictive (P) slice or picture a single respective reference picture is employed for encoding and decoding each block of pixels in the P slice or picture. A bi-predictive (B) slice or picture can use two respective reference pictures for encoding and decoding each block of pixels in the B slice or picture. Hence, such P and B slices and pictures are encoded and decoded with reference to and depending on other pictures, denoted reference pictures, in the video sequence.

Hence, if a picture B is encoded based on data in another picture A in the video sequence this picture B directly depends on the picture A. It could also be possible that a further picture C is encoded based

on data in the picture B. This means that this picture C indirectly depends on the picture A since picture C is encoded relative to the picture B, which in turn is encoded relative to the picture A.

In a particular embodiment of the method in Fig. 4, a list of multiple identifiers, such as access unit identifiers or picture identifiers, is provided in step S1. In such a case, in an embodiment step S2 is performed for each access unit identified from the list. For instance, if the list identifies N access units, step S2 is preferably performed N times, i.e. one for each identified access unit.

Hence, in an embodiment a list of values is given as input to the sub-bitstream extraction process. The access units that correspond to these values are removed from the bitstream. Further, all access units that depend on/can depend on (reference/can reference) the removed access units are removed from the sequence, i.e. video sequence, iteratively until there are no access unit in the resulting sub-bitstream that depends on/can depend on (references/can reference) a removed access unit.

Each picture of the video sequence preferably has, in an embodiment, a respective layer identifier identifying a respective layer in the video sequence. In this embodiment step S2 preferably comprises discarding the access unit identified in step S1 and all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to a decoding order of the bitstream but preceding, according to a defined order of the bitstream, any next layer switching point having a layer identifier lower than or equal to the layer identifier of the access unit identified in step S1.

The above disclosed embodiment therefore relates to so-called layered video sequences wherein the video sequence comprises multiple, i.e. at least two, layers of pictures. Such layered video sequences are typically employed, for instance, to achieve scalability, including temporal scalability, spatial scalability, etc., and/or in the case of multi-view video, such as stereoscopic video. In such a case, each picture could have a respective temporal identifier and/or view identifier as illustrative examples, where the temporal identifier identifies the temporal layer of the video sequence to which the picture belongs and the view identifier identifies the camera view to which the picture belongs to.

30

A layer switching point is a point in the bitstream at which it is possible to switch decoding of more layers of the video sequence than what was decoded before the layer switching. Hence, the layer switching points enables a decoder, during the decoding process of the bitstream, to start decoding

and thereby outputting, such as for display, pictures belonging to a higher layer than what is currently being decoded and output.

There are generally two variants of layer switching points. In a first variant, also denoted “switch-from”
5 variant, the layer switching point indicates the picture where it is possible to switch from a current layer to decoding more layers of the video sequence. In a second variant, also denoted “switch to” variant, the layer switching point indicates the picture that it is possible to do layer switching to.

In an embodiment, only the first or “switch-from” variant is available, whereas in other embodiments
10 only the second or “switch-to” variant is available or both the first and second variants are available.

If the next layer switching point is of the “switch-from” variant step S2 preferably comprises discarding all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to the
15 decoding order but preceding, according to the defined order, a next layer switching point (of the “switch-from” variant) having a layer identifier lower than the layer identifier of the access unit identified in step S1.

Correspondingly, if the next layer switching point is of the “switch-to” variant step S2 preferably
20 comprises discarding all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to the decoding order but preceding, according to the defined order, a next layer switching point (of the “switch-to” variant) having a layer identifier lower than or equal to the layer identifier of the access unit identified in step S1.

25

In a particular embodiment, each picture of the video sequence has a respective temporal identifier, *temporal_id*, identifying a respective temporal layer of the video sequence. In such a case, a layer switching point is a temporal layer switching point. Such a temporal layer switching point is preferably selected from a group consisting of a Temporal Sub-layer Access (TSA) picture, also denoted
30 Temporal Layer Access (TLA) picture; an Instantaneous Decoding Refresh (IDR) picture and other Random Access Point (RAP) pictures.

The embodiment of step S2 that is adapted to layered video compares the layer identifier of following, according to the decoding order, access units in the bitstream with the layer identifier of the access unit

identified in step S1. In more detail, if a following, according to the decoding order, access unit carries a picture with a layer identifier that is at least equal to the layer identifier of the access unit identified in step S1 this access unit is identified as an access unit that can depend on the access unit identified in step S1. This is possible since, in an embodiment, a lower layer should never depend on a higher layer. Thus, a picture belonging to a layer L_n could be encoded depending on or referencing a picture in layer L_n or in layer L_{n-m} but not relative to a picture in layer L_{n+k} , wherein k, m, n are positive integers and $n > m$.

In this embodiment the layer switching points are used to detect the limits of which pictures and access units that need to be removed. Hence, in step S2 preferably only access units preceding, according to the defined order of the bitstream, any next layer switching point are considered with regard to discarding to form the sub-bitstream. This is possible since it is guaranteed that a picture that follows a layer switching point is not affected by, i.e. does not depend on, a picture of the same or higher temporal layer as the switching point, which precedes the layer switching point.

15

A particular embodiment therefore uses three criteria in order to determine or identify an access unit in the bitstream as an access unit that can depend on the access unit identified in step S1:

- 1) the access unit has a layer identifier, such as temporal identifier, that is higher than or equal to the layer identifier, such as temporal identifier, of the access unit identified in step S1;
- 20 2) the access unit follows the access unit identified in step S1 according to the decoding order of the bitstream; and
- 3) the access unit precedes, according to the defined order of the bitstream, any next layer switching point, such as temporal layer switching point, having a layer identifier, such as temporal identifier, lower than or equal ("switch-to" variant) or lower than ("switch from" variant) to the layer identifier, such as temporal identifier, of the access unit identified in step S1.

25

HEVC also defines a Stepwise Temporal Sub-layer Access point (STSA). For this type of switching point the indication has to have a layer identifier, such as temporal identifier, exactly equal to the layer identifier, such as temporal identifier, of the layer that you would like to switch to in order to make switching possible. Thus, in such a case the criterion 3) is:

30

- 3) the access unit precedes, according to the defined order of the bitstream, any next STSA having a temporal identifier equal the temporal identifier of the access unit identified in step S1.

In an embodiment, temporal layer switching points with certain values of *temporal_id* are used to find the limits for which pictures can depend on a removed picture. By using temporal layer switching points to find the limits, the removal of dependent pictures only needs to be applied once, i.e. not in an iterative fashion, since there is no dependency across a temporal layer switching point of a certain

5 *temporal_id*.

The defined order of the bitstream mentioned above in criterion 3) is, in an embodiment, the decoding order of the access units in the bitstream. In another embodiment, the defined order of the bitstream is the output order of the decoded pictures represented by the encoded pictures in the bitstream.

10

The decoding order defines the order at which coded pictures of the bitstream are decoded. Hence, a picture with decoding order number N is decoded prior to a picture with decoding order number N+1. Output order defines the order at which decoded pictures are output. Output of pictures typically implies output for display on a screen of or connected (wired or wirelessly) to the decoder. Output of pictures

15 could, however, be for other purposes than display, for instance, output for storage in a file, output for transcoding, or output for post-decoding processing.

Generally, a picture with output order number M is output prior to a picture with output order number M+1. Note that a picture with decoding order number N+k could be output prior to a picture with

20 decoding order number N, where k is a positive integer. In HEVC output order is specified by a Picture Order Count (POC) value.

An embodiment of step S2 involves discarding the access unit identified in step S1 from the bitstream. Step S2 additionally comprises discarding all access units in the bitstream meeting the criteria 1)-3)

25 above if the bitstream comprises a next layer switching point having a layer identifier lower than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to the defined order in the bitstream. If the bitstream does not comprise any such next layer switching point having a layer identifier lower than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to the

30 defined order step S2 instead preferably comprises discarding all access units in the bitstream meeting the criteria 1) and 2) above.

Thus, in the former case all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit identified in step S1 and following the access unit according to the

decoding order but preceding, according to the defined order, a next layer switching point having a layer identifier lower than or equal to the layer identifier of the access unit identified in step S1 are discarded if the bitstream comprises such a next layer switching point. If no such next layer switching point is present in the bitstream step S2 preferably comprises discarding all access units in the
5 bitstream having layer identifier higher than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to the decoding order.

In a particular embodiment step S2 comprises discarding the access unit identified in step S1 and all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the
10 access unit identified in step S1 and following the access unit identified in step S1 according to a decoding order of the bitstream but preceding, according to the decoding order, any next layer switching point having a layer identifier lower than or equal to the layer identifier of the access unit identified in step S1.

15 In another particular embodiment step S2 comprises discarding the access unit identified in step S1 and all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit identified in step S1 and following the access unit identified in step S1 according to a decoding order of the bitstream but preceding, according to an output order of the bitstream, any next layer switching point having a layer identifier lower than or equal to the layer identifier of the access unit
20 identified in step S1.

The discarding in step S2 of the access unit identified in step S1 and of all the access units that can depend on this access unit is preferably performed without decoding this access unit. This significantly reduces the complexity of the sub-bitstream extracting method as compared to first decoding the
25 access unit and the coded picture contained therein and then discarding the decoded picture and access unit.

Fig. 5 is a flow diagram illustrating an additional, optional step of the method in Fig. 4. The method continues from step S1 in Fig. 4. In a next step S10 the access unit identified in step S1 is marked as
30 "to be removed from the bitstream". This step S10 preferably also comprises marking all access units in the bitstream that can depend on the access unit identified in step S1 as "to be removed from the bitstream". The method then continues to step S2 of Fig. 4. In this embodiment step S2 preferably comprises discarding all access units of the bitstream that are marked as "to be removed from the

bitstream". Hence, the access unit or picture marking is preferably employed in order to identify those access units or pictures of the bitstream that should be discarded.

Marking an access unit as "to be removed from the bitstream" could be performed by marking all NAL units of that access unit as "to be removed from the bitstream". In such a case, step S2 preferably comprises removing or discarding all NAL units that are marked as "to be removed from the bitstream".

A particular implementation example will now be further described with regard to the sub-bitstream extraction process.

10

Input to this process is an ordered list of integers *au_num* (access unit number), in increasing order, with N elements, where N is in the range of 0 to the number of access units in the bitstream, inclusive. Each element *au_num[i]* in the list, for i in the range from 0 to N-1, has a value in the range from 0 to the number of access units in the bitstream.

15

Output of this process is a sub-bitstream that conforms to the video coding standard or specification.

The sub-bitstream is derived by applying the following operations in sequential order:

20 1. For each i in the range of 0 to N-1 the following applies:

- a. Let $j = au_num[i]$, let auX be the j^{th} access unit in decoding order and, when there is an access unit with a TLA picture or IDR picture with *temporal_id* lower than or equal to the *temporal_id* of auX that follows auX in output order, call that access unit auY.

25 In an embodiment, auY may precede auX in decoding order.

- b. If there is an access unit auY, then for each access unit auZ with *temporal_id* higher than or equal to the *temporal_id* of auX that follows auX in the decoding order and precedes auY in output order, mark all NAL units of auZ as "to be removed from the bitstream".

30

- c. Otherwise, i.e. there is no access unit auY, then for each access unit auZ with *temporal_id* higher than or equal to the *temporal_id* of auX that follows auX in decoding order, mark all NAL units of auZ as "to be removed from the bitstream".

2. Remove all NAL units marked as “to be removed from the bitstream”.

In another implementation example the sub-bitstream is derived by applying the following operations in sequential order:

5

1. For each i in the range of 0 to $N-1$ the following applies:
 - a. Let $j = au_num[i]$, let auX be the j^{th} access unit in decoding order and, when there is an access unit with a TLA picture or IDR picture with *temporal_id* lower than or equal to the *temporal_id* of auX that follows auX in decoding order, call that access unit auY .

10

- b. If there is an access unit auY , then for each access unit auZ with *temporal_id* higher than or equal to the *temporal_id* of auX that follows auX in the decoding order and precedes auY in the decoding order, mark all NAL units of auZ as “to be removed from the bitstream”.

- c. Otherwise, i.e. there is no access unit auY , then for each access unit auZ with *temporal_id* higher than or equal to the *temporal_id* of auX that follows auX in decoding order, mark all NAL units of auZ as “to be removed from the bitstream”.

2. Remove all NAL units marked as “to be removed from the bitstream”.

20

In a particular embodiment of these implementation examples, when *au_num* contains the integers that correspond to all access units with *temporal_id* higher than a variable *tIdTarget*, the entire sequence will be scaled to only contain temporal layers lower than or equal to *tIdTarget*.

- 25 The sub-bitstream extraction process can be performed in an encoder, a decoder or in a network node.

In an embodiment, when a “switch-to” variant or definition is used for switching points a decoder may use the sub-bitstream extraction process as exemplified below.

- 30 When decoding a bitstream a decoder may choose to discard an access unit without decoding it, e.g. for complexity reasons.

By also discarding all access units with the same or higher *temporal_id* as the discarded picture that precedes the next temporal layer switching point, e.g. TLA picture, in output order (or alternatively

decoding order) with a *temporal_id* lower than or equal to the *temporal_id* of the discarded picture, the decoder can still decode all other access units and be guaranteed that it is a valid bitstream because the pictures that follow the temporal layer switching point are not in any way affected by the discarded pictures. If there is no temporal layer switching point that follows the discarded picture with a
5 *temporal_id* lower than or equal to the *temporal_id* of the discarded picture, all access units with the same or higher *temporal_id* as the discarded picture are discarded until the end of the bitstream.

In an embodiment, when a “switch-from” variant or definition is used for switching points a decoder may use the sub-bitstream extraction process as exemplified below.

10

When decoding a bitstream a decoder may choose to discard an access unit without decoding it, e.g. for complexity reasons.

By also discarding all access units with the same or higher *temporal_id* as the discarded picture that
15 precedes the next temporal layer switching point in output order (or alternatively decoding order) with a *temporal_id* lower than the *temporal_id* of the discarded picture, the decoder can still decode all other access units and be guaranteed that it is a valid bitstream because the pictures that follow the temporal layer switching point are not in any way affected by the discarded pictures. If there is no temporal layer switching point that follows the discarded picture with a *temporal_id* lower than the *temporal_id* of the
20 discarded picture, all access units with the same or higher *temporal_id* as the discarded picture are discarded until the end of the bitstream.

An advantage of embodiments is that each sub-bitstream created from a conforming bitstream using the sub-bitstream extraction process disclosed herein is a conforming bitstream. Hence, no bitstream
25 conformance verification or check needs to be conducted on the formed sub-bitstream.

In an optional embodiment a condition may be added for bitstream conformance in order to ensure that the Decoded Picture Buffer (DPB) requirements are fulfilled for each temporal layer, see condition 6 in Annex A and condition 5 in Annex B.

30

At any time instant, there shall be at most *max_dec_frame_buffering[i]* pictures with *temporal_id* lower than or equal to *i* in the DPB for all *i* in the range of 0 to *max_temporal_layers_minus1*, inclusive. The parameter or syntax element *max_dec_frame_buffering[i]* is preferably signaled for each temporal layer, preferably in the SPS or in another parameter set.

In an optional embodiment, the condition mentioned above replaces the condition that the fullness of the DPB shall be less than or equal to the DPB size, see condition 6 in Annex A and condition 5 in Annex B.

5

A further advantage of the embodiments is that sub-bitstream extraction process can remove or discard access units and pictures from anywhere in the bitstream. Hence, the sub-bitstream extraction process does not require the removal of complete layers, i.e. scaling the entire video sequence to a specific layer. The sub-bitstream extraction process does furthermore not require the removal of entire
10 layers between layer switching points. Hence, the removal of pictures can start anywhere in the bitstream and does not need to be aligned with or started from such a layer switching point. As a consequence, a very flexibly but also computationally efficient sub-bitstream process is achieved.

In a particular embodiment, the access unit identified in step S1 of Fig. 4 is not a layer switching point,
15 such as TLA picture. Hence, in a particular embodiment, the access unit preferably comprises a coded picture that is present in the bitstream between such layer switching points.

The present embodiments are in particular suitable for usage in connection with HEVC. In such a case, the sub-bitstream extraction method and device are applicable to a HEVC-conforming bitstream of
20 multiple access units.

In a particular implementation example, the embodiments are applied to HEVC-conforming bitstreams of multiple access units and HEVC codecs. In this implementation example, the bistream comprises the previously mentioned "switched to" and/or "switch from" TLA pictures. Furthermore, a Constant Bit Rate
25 (CBR) flag, *cbr_flag*, is preferably set to 0. In such a case, a *cbr_flag* set to 0 preferably specifies that to decode the bitstream by a Hypothetical Reference Decoder (HRD), a Hypothetical Stream Delivery Scheduler (HSS) operates in an intermittent bit rate mode.

HRD is a hypothetical decoder model that specifies constraints on the variability of conforming NAL unit
30 streams or conforming byte streams that an encoding process may produce. A HSS is a hypothetical delivery mechanism for the timing and data flow of the input of a bitstream into the HRD. The HSS is used for checking the conformance of a bitstream or a decoder.

An aspect of the embodiments relates to a method in an encoder for bitstream extraction, i.e. the sub-bitstream extraction method can be performed in an encoder.

Another aspect of the embodiments relates to a method in a decoder for bitstream extraction, i.e. the
5 sub-bitstream extraction method can be performed in a decoder.

Fig. 6 is a schematic block diagram of a sub-bitstream extracting device 30 according to an embodiment. The sub-bitstream extracting device 30 preferably comprises an access unit identifier 31, also denoted access unit identifying unit, means or module. The access unit identifier 31 is configured
10 to identify an access unit among a bitstream of multiple access units. Each such access unit of the bitstream comprises a respective encoded picture of a video sequence. An access unit discarder 32, also denoted access unit discarding unit, means or module, is configured to discard the access unit identified by the access unit identifier 31 and all access units in the bitstream that can depend on the access unit. The operation of the access unit discarder 32 forms a sub-bitstream having no access
15 units that can depend on the access unit as identified by the access unit identifier 31.

In a particular embodiment the access unit identifier 31 is further configured to identify the access units in the bitstream that can depend on the access unit. This can, for instance, be performed by identifying those access units in the bitstream that directly or indirectly depends on or references the particular
20 access unit as disclosed herein. In a particular embodiment, the access unit identifier 31 is configured to identify the access units in the bitstream that can depend on the particular access unit based on respective layer identifiers of the access units and the layer identifier of the particular access unit. Furthermore, the access unit identifier 31 preferably also uses respective access unit or picture numbers in the identification of the access units in order to perform the identification among access
25 units that follow the particular access unit according to the decoding order.

Thus, in an embodiment each picture of the video sequence has a respective layer identifier identifying a respective layer. The access unit discarder 32 is then configured to discard the access unit and all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the
30 access unit and following this access unit according to the decoding order but preceding, according to a defined order, such as output order or decoding order, any next layer switching point having a layer identifier lower than or equal to the layer identifier of this access unit.

If the next layer switching point is of the previously mentioned “switch-to” variant, the relevant next layer switching point preferably has a layer identifier, such as *temporal_id*, lower than or equal to the layer identifier, such as *temporal_id*, of the access unit identified by the access unit identifier 31.

- 5 If the next layer switching point is instead of the previously mentioned “switch-from” variant, the relevant next layer switching point preferably has a layer identifier, such as *temporal_id*, lower than the layer identifier, such as *temporal_id*, of the access unit identified by the access unit identifier 31.

In a particular embodiment the access unit discarder 32 is configured to discard the access unit
10 identified by the access unit identifier 31 from the bitstream. The access unit discarder 32 is also configured to discard all access units in the bitstream having layer identifier higher than or equal to the layer identifier of this access unit and following this access unit according to the decoding order but preceding, according to the defined order, a next layer switching point having a layer identifier lower than or equal to the layer identifier of this access unit if the bitstream comprises such a layer switching
15 point. If there are no such layer switching point in the bitstream the access unit discarder 32 is preferably configured to discard all access units in the bitstream having layer identifier higher than or equal to the layer identifier of this access unit and following this access unit according to the decoding order.

- 20 In a first implementation embodiment the access unit discarder 32 is configured to discard the access unit and all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit and following this access unit according to the decoding order but preceding, according to the decoding order, any next layer switching point having a layer identifier lower than or equal to the layer identifier of this access unit.

25

In a second implementation embodiment the access unit discarder 32 is configured to discard the access unit and all access units in the bitstream having layer identifier higher than or equal to the layer identifier of the access unit and following this access unit according to the decoding order but preceding, according to the output order, any next layer switching point having a layer identifier lower
30 than or equal to the layer identifier of this access unit.

The sub-bitstream extracting device 30 optionally comprises a marking unit 33, also denoted access unit marker or marking means or module. The marking unit 33 is configured to mark the access unit identified by the access unit identifier 31 as “to be removed from the bitstream”. The marking unit 33 is

preferably also configured to mark all access units in the bitstream that can depend on this access unit as “to be removed from the bitstream”.

The access unit discarder 32 is, in this embodiment, preferably configured to discard or remove all
5 access units of the bitstream that are marked as “to be removed from the bitstream”.

As previously disclosed herein the access unit discarder 32 is preferably configured to discard the access units prior to decoding them.

10 The sub-bitstream extracting device 30 of Fig. 6 with its including units 31-33 could be implemented in hardware. There are numerous variants of circuitry elements that can be used and combined to achieve the functions of the units 31-33 of the sub-bitstream extracting device 30. Such variants are encompassed by the embodiments. Particular examples of hardware implementation of the sub-bitstream extracting device
30 is implementation in digital signal processor (DSP) hardware and integrated circuit technology,
15 including both general-purpose electronic circuitry and application-specific circuitry.

The sub-bitstream extracting device 30 can also be implemented by means of a processor and a memory. Thus, in an embodiment the sub-bitstream extracting device 30 is implemented e.g. by one or more of a processor and adequate software with suitable storage or memory therefore, a
20 programmable logic device (PLD) or other electronic component(s).

An aspect of the embodiments relates to a decoder for bitstream extraction. Fig. 7 is a schematic block diagram of such a decoder 40 comprising a sub-bitstream extracting device 30 as described above and disclosed in Fig. 6. In an alternative embodiment, the units of the sub-bitstream extracting device
25 30, i.e. the access unit identifier 31, the access unit discarder 32 and the optional marking unit 33 could be directly implemented in the decoder 40.

The functionality of the decoder 50 can be implemented by means of a processor 52 and a memory 54 as illustrated in Fig. 8.

30

In addition, the decoder 50 preferably comprises an input or input unit 51 configured to receive the access units or coded pictures of the bitstreams. A corresponding output or output unit 53 is configured to output decoded pictures.

The decoder can be implemented in a device, such as a mobile device exemplified as mobile phones, tablets, video camera, set-top-box, etc. Fig. 9 illustrates such an example where the decoder 62 is located in a receiver 60. The receiver 60 then comprises an input or input unit 61 configured to receive a coded bitstream. The coded bitstream is preferably, in addition to being decoded, also subject to a sub-bitstream extraction process to thereby, for at least a portion of the bitstream, form such a sub-bitstream that could be decoded by the decoder 62. The decoder 62 preferably comprises or is connected to a reference picture buffer 64, i.e. DPB, that temporarily stores already decoded reference pictures 65 that are to be used as reference pictures for other pictures in the video stream. Decoded pictures are output from the receiver 60, such as from the reference picture buffer 64, by means of an output or output unit 63. These output pictures are typically sent to be displayed to a user on a screen or display of or connected, including wirelessly connected, to the receiver 60.

An aspect of the embodiments relates to an encoder for bitstream extraction. Fig. 10 is a schematic block diagram of such an encoder 70 comprising a sub-bitstream extracting device 30 as described above and disclosed in Fig. 6. In an alternative embodiment, the units of the sub-bitstream extracting device 30, i.e. the access unit identifier 31, the access unit discarder 32 and the optional marking unit 33 could be directly implemented in the encoder 70.

An encoder may use the embodiments by setting the layer values, such as *temporal_id* values, of the access units in the bitstream and by only validating the conformance of the full bitstream. Hence, there is no need to validate conformance of the sub-bitstreams as it follows by the sub-bitstream extraction process definition that they will be conforming bitstreams. Thus, the conformance of the sub-bitstreams is guaranteed by the standard.

The encoder can signal layer switching points, such as temporal layer switching points, for each picture for which the reference are according to the definition of the layer switching point.

The encoder does not need to consider the DPB status for all different possible extracted sub-bitstreams.

30

The functionality of the decoder 80 can be implemented by means of a processor 82 and a memory 84 as illustrated in Fig. 11.

In addition, the encoder 80 preferably comprises an input or input unit 81 configured to receive the pictures of the video sequence. A corresponding output or output unit 83 is configured to output the coded pictures, such as in the form of access units comprising NAL units.

- 5 The encoder can be implemented in a device, such as a mobile device exemplified as mobile phones, tablets, video camera, etc. Fig. 12 illustrates an example of such a device in the form of a transmitter 90. The transmitter 90 then comprises an input or input unit 91 configured to receive pictures of a video sequence to be encoded. The pictures are encoded by the encoder 92 to form a coded bitstream of access units. The coded bitstream is output from the transmitter 90 by an output or output unit 93. The
10 encoder 92 is preferably configured to form a sub-bitstream as disclosed herein from the bitstream.

Another aspect of the embodiments relates to a method in a network node for bitstream extraction. The network node can start discarding access units with a layer identifier, such as *temporal_id*, higher than a selected layer identifier X from a relayed bitstream at any point in time. The network node can then
15 start relaying access units with layer identifier, such as *temporal_id*, higher than X at any layer switching point, such as TLA picture, with layer identifier, such as *temporal_id*, equal to X or X+1.

The advantage of this sub-bitstream extraction method is the discarding of access units and the relaying of access units can be performed without checking the resulting sub-bitstream for bitstream
20 conformance.

Fig. 13 is a flow diagram illustrating an embodiment of such a sub-bitstream extraction method. The sub-bitstream extraction method is configured to be performed in or by a network node. The method starts in step S20, in which access units of a bitstream of multiple access units are discarded to form a
25 sub-bitstream. Each access unit of the bitstream comprises a respective encoded picture of a video sequence. Each picture of the video sequence has a respective layer identifier identifying a respective layer to which the picture belongs. In step S20 access units with layer identifier higher than a selected layer identifier X are discarded from the bitstream to form the sub-bitstream.

- 30 A next step S21 comprises relaying access units with layer identifier higher than X starting from a layer switching point with a layer identifier equal to X or X+1 to form an updated bitstream.

The discarding of access units in step S20 and the relaying of access units in step S21 are performed without checking the sub-bitstream and the updated bitstream for bitstream conformance.

The network node preferably receives the bitstream originating from a transmitter as disclosed in Fig. 12 and should relay or forward the access units of the bitstream towards a receiver as disclosed in Fig. 9.

5

In an embodiment of step S20, the discarding of access unit is performed based on information specifying that access units belonging to layers higher than the selected layer identifier X should be removed from the bitstream. This information specifying the value X could be received from an external unit, such as a transmitter of Fig. 12, an encoder of Figs. 10 or 11, a receiver of Fig. 9 or a decoder of
10 Figs. 7 or 8. Alternatively, the network node itself could determine this value X based on various input parameters, such as current available bandwidth for the bitstream and/or a desired current bitrate for the bitstream.

The discarding of step S20 thereby removes access units from one or more layers starting at a
15 selected point in the bitstream. The access units of the bitstream that are not removed from this selected point form the sub-bitstream. These access units of the sub-bitstream are preferably relayed or forwarded by the network node towards an intended receiver and decoder.

At a later selected point in the bitstream the network node should start relaying access units belong to
20 a higher layer than X. Thus, at this later selected point more layers of pictures should be relayed and forwarded towards the receiver and decoder than prior to the later selected point.

For instance, assume a video sequence comprises pictures of four different layers denoted layer 0 to 3. Step S20 could then, as an illustrative but non-limiting example, discard access units belonging to layer
25 2 and access units belonging to layer 3. Hence, the sub-bitstream that is relayed by the network node then carries access units with coded pictures belonging to layer 0 or layer 1. The network node can then start relaying pictures of a higher layer to form the updated bitstream. For instance, this updated bitstream could comprise access units with coded pictures belonging to layer 0-2 or indeed access units with coded pictures belonging to layer 0-3.

30

The selection of which layers to include in the updated bitstream could be performed based on information as received from the transmitter or receiver, or indeed based on updated bandwidth and/or bitrate parameters.

The relaying of the updated bitstream occurs at a layer switching point that either has a layer identifier equal to X or equal to X+1, which is further discussed herein.

A significant advantage of the sub-bitstream extraction method disclosed above as compared to prior art extraction processes is that the discarding and relaying of access units in steps S20 and S21 can be performed without the need for checking or verifying that the formed sub-bitstream and updated bitstream are conforming bitstreams. In clear contrast, by performing the discarding and relaying as disclosed herein the sub-bitstream and the updated bitstream are guaranteed to be conforming bitstreams.

10

Checking the sub-bitstream and the updated bitstream for bitstream conformance corresponds to verifying that the sub-bitstream and the updated bitstream fulfill the requirements of the relevant video coding standard. In particular, checking for bitstream conformance comprises verifying that each of the conditions specified in Annex A for HEVC or in Annex B for H.264/AVC are fulfilled or other such conditions as defined for the relevant video coding standard.

In an embodiment, each picture of the video sequence has a respective temporal identifier, *temporal_id*, identifying a respective temporal layer. In an embodiment, the layer switching point is a TLA picture, to which it is possible to perform temporal layer switching, i.e. the so-called switch-to variant. In such a case, step S21 preferably comprises relaying, to form the updated bitstream, access units with temporal identifier higher than X starting from the TLA picture with a temporal identifier equal to X+1 in the bitstream.

Thus, when a switch-to definition is used for switching points a network node may perform the following ordered steps. The network node can start discarding access units with *temporal_id* higher than X from a relayed bitstream at any point in time, for any value of X as selected by the network node. It can then start relaying access units with *temporal_id* higher than X at any temporal layer switching point, e.g. TLA picture, with *temporal_id* equal to X+1 because these access units are not in any way affected by the discarded picture.

30

In another embodiment, the so-called switch-from variant is available, i.e. the layer switching point is a TLA picture from which it is possible to perform temporal layer switching. Step S21 preferably comprises relaying access units, to form the updated bitstream, with temporal identifier higher than X starting from the TLA picture with a temporal identifier equal to X in the bitstream.

When a switch-from definition is used for switching points a network node may perform the following ordered steps. The network node can start discarding access units with *temporal_id* higher than X from a relayed bitstream at any point in time for any value of X as selected by the network node. It can then
5 start relaying access units with *temporal_id* higher than X at any temporal layer switching point, e.g. TLA picture, with *temporal_id* equal to X because these access units are not in any way affected by the discarded pictures.

In a particular embodiment the method as illustrated in Fig. 13 may optionally comprise declaring that
10 the formed sub-bitstream and the updated bitstream are conforming bitstreams and where this declaration is performed without checking the sub-bitstream and the updated bitstream for bitstream conformance.

The optional declaration may in a particular example be in the form of setting a flag or other syntax
15 element to a defined value indicating that the formed sub-bitstream and the updated bitstream are conforming bitstreams.

Fig. 14 is a schematic block diagram of an embodiment of a network node 80. The network node 80 can be any device or unit that is arranged in or connected to a communication network, including
20 mobile or cellular networks, in which coded bitstreams are communicated between a transmitter or encoder and a receiver or decoder. The network node 80 could for instance represent a base station, also denoted base radio station or node-B in the art. Alternatively, the network node 80 could be a device implemented in or connected (wired or wirelessly) to such a base station. The network node 80 could, for instance, form part of or constitute a media processing device implemented in the
25 communication network.

The network node 80 generally comprises an access unit discarder 81, also denoted access unit discarding unit, means or module. The access unit discarder 81 is configured to discard access units of a bitstream of multiple access units comprising a respective encoded picture of a video sequence. As
30 previously mentioned herein, each picture of the video sequence has a respective layer identifier identifying a respective layer. The access unit discarder 81 is then configured to discard access units with layer identifier higher than a selected layer identifier X from the bitstream to form a sub-bitstream.

The value X of the selected layer identifier could be determined by the network node 80 itself, such as by a processor (not illustrated) of the network node 80 as previously described herein. Alternatively, the network node 80 receives a notification of this value X from an external unit in the communication network.

5

The network node 80 also comprises a relaying unit 82, also denoted access unit relayer or relaying means or module. The relaying unit 82 is configured to relay, to form an updated bitstream, access units of the bitstream with layer identifier higher than X starting from a layer switching point with a layer identifier equal to X or X+1 in the bitstream.

10

The access unit discarder 81 and the relaying unit 82 are operated for discarding the access units and relaying the access units, respectively, without any need for checking the sub-bitstream and the updated bitstream for bitstream conformance.

15 In an embodiment, the network node 80 also comprises an input or input unit 83 that is configured to receive the coded bitstream with its access units. This input unit 83 may optionally also receive a notification of the selected value X for the layer identifier. A corresponding optional output or output unit 84 is preferably arranged in the network node 80 to output the coded sub-bitstream and the updated bitstream.

20

If each picture of the video sequence has a respective temporal identifier and the layer switching point is a TLA picture to which it is possible to perform a temporal layer switching then the relaying unit 82 is preferably configured to relay, to form the updated bitstream, the access units with temporal identifier higher than X starting from the TLA picture with a temporal identifier equal to X+1 in the bitstream.

25

If the layer switching point is instead a TLA picture from which it is possible to perform a temporal layer switching then the relaying unit 82 is preferably configured to relay, to form the updated bitstream, the access units with temporal identifier higher than X starting from the TLA picture with a temporal identifier equal to X in the bitstream.

30

Thus, an aspect of the embodiments relates to a network node 80 for bitstream extraction. The network node 80 can start discarding access units with *temporal_id* higher than X from a relayed bitstream at any point in time. The network node 80 can then start relaying access units with *temporal_id* higher than X at any TLA picture with *temporal_id* equal to X+1 or equal to X.

In a particular embodiment the network node 80 is configured to declare that the formed sub-bitstream and the updated bitstream are conforming bitstreams and where this declaration is performed without checking the sub-bitstream and the updated bitstream for bitstream conformance.

5

The network node 80 of Fig. 14 with its including units 81-84 could be implemented in hardware. There are numerous variants of circuitry elements that can be used and combined to achieve the functions of the units 81-84 of the network node 80. Such variants are encompassed by the embodiments. Particular examples of hardware implementation of the network node 80 are implementation in DSP hardware and
10 integrated circuit technology, including both general-purpose electronic circuitry and application-specific circuitry.

The network node 80 can also be implemented by means of a processor and a memory. Thus, in an embodiment the network node 80 is implemented e.g. by one or more of a processor and adequate
15 software with suitable storage or memory therefore, a PLD or other electronic component(s).

The embodiments can be applied to different video codecs and different types of extensions including, but not limited to, multi-view video codecs and scalable video codecs. For such extension the temporal layers can be replaced by other layers including, but not limited to, spatial layers, quality layers,
20 dependency layers or different views. A switching point would for those cases be a spatial layer switching point, a quality layer switching point, a dependency layer switching point and a view switching point.

The embodiments described above are to be understood as a few illustrative examples of the present
25 invention. It will be understood by those skilled in the art that various modifications, combinations and changes may be made to the embodiments without departing from the scope of the present invention. In particular, different part solutions in the different embodiments can be combined in other configurations, where technically possible. The scope of the present invention is, however, defined by the appended claims.

30

ANNEX A

The present Annex A lists the ten conditions specified for HEVC in order to achieve bitstream conformance. In order to a bitstream of coded data to be regarded as meeting the requirement of bitstream conformance all of the following conditions should be fulfilled.

1. For each access unit n , with $n > 0$, associated with a buffering period Supplemental Enhancement Information (SEI) message, with $\Delta t_{g,90}(n)$ specified by

$$\Delta t_{g,90}(n) = 90000 \times (t_{r,n}(n) - t_{af}(n-1))$$

5 the value of *InitCpbRemovalDelay*[*SchedSelIdx*] shall be constrained as follows.

If *cbr_flag*[*SchedSelIdx*] is equal to 0,

$$\text{InitCpbRemovalDelay}[\text{SchedSelIdx}] \leq \text{Ceil}(\Delta t_{g,90}(n))$$

Otherwise (*cbr_flag*[*SchedSelIdx*] is equal to 1),

$$\text{Floor}(\Delta t_{g,90}(n)) \leq \text{InitCpbRemovalDelay}[\text{SchedSelIdx}] \leq \text{Ceil}(\Delta t_{g,90}(n))$$

10 The exact number of bits in the Coded Picture Buffer (CPB) at the removal time of each picture may depend on which buffering period SEI message is selected to initialize the Hypothetical Reference Decoder (HRD). Encoders must take this into account to ensure that all specified constraints must be obeyed regardless of which buffering period SEI message is selected to initialize the HRD, as the HRD may be initialized at any one of the buffering period SEI messages.

15

CPB is a first-in first-out buffer containing access units in decoding order specified in the HRD. The HRD is a hypothetical decoder model that specifies constraints on the variability of conforming NAL unit streams or conforming byte streams that an encoding process may produce.

20 2. A CPB overflow is specified as the condition in which the total number of bits in the CPB is larger than the CPB size. The CPB shall never overflow.

3. A CPB underflow is specified as the condition in which the nominal CPB removal time of decoding unit m $t_{r,n}(m)$ is less than the final CPB arrival time of decoding unit m $t_{af}(m)$ for any value of m . When

25 *low_delay_hrd_flag* is equal to 0, the CPB shall never underflow.

4. When *low_delay_hrd_flag* is equal to 1, a CPB underflow may occur at decoding unit m . In this case, the final CPB arrival time of access unit n containing decoding unit m $t_{af}(n)$ shall be later than the nominal CPB removal time of access unit n containing decoding unit m $t_{r,n}(n)$.

30

5. The nominal removal times of pictures from the CPB (starting from the second picture in decoding order), shall satisfy the constraints on $t_{r,n}(n)$ and $t_r(n)$ expressed in subclauses A.4.1 - General tier and level limits and A.4.2 - Profile-specific level limits for the Main and Main 10 profiles of the HEVC specification.

6. After the decoding process for reference picture set as specified in subclause 8.3.2 - Decoding process for reference picture set of the HEVC specification has been invoked, the number of decoded pictures in the Decoded Picture Buffer (DPB), including all pictures marked as "used for reference" and
 5 all pictures n that have $PicOutputFlag$ equal to 1 and $t_c(n) < t_c(currPic)$ and $t_{o,dpb}(n) \geq tr(currPic)$, where $currPic$ is the current picture, but not including the current picture, with $TemporalId$ lower than or equal to the $TemporalId$ of the current picture, shall be less than or equal to $\text{Min}(0, sps_max_dec_pic_buffering[TemporalId] - 1)$.

10 The DPB is a buffer holding decoded pictures for reference, output reordering, or output delay specified for the HRD.

7. All reference pictures shall be present in the DPB when needed for prediction. Each picture shall be present in the DPB at its DPB output time unless it is removed from the DPB before its output time by
 15 one of the processes specified in subclause C.3 - Operation of the decoded picture buffer (DPB) of the HEVC specification.

8. The value of $maxPicOrderCnt - minPicOrderCnt$ shall be less than $MaxPicOrderCntLsb / 2$.

20 9. The value of $\Delta t_{o,dpb}(n) = t_{o,dpb}(n_n) - t_{o,dpb}(n)$, which is the difference between the output time of a picture and that of the first picture following it in output order and having pic_output_flag equal to 1, shall satisfy the constraint expressed in subclause General tier and level limits of the HEVC specification for the profile, tier and level specified in the bitstream using the decoding process specified in clauses 2 - Normative references; 3 - Definitions; 4 - Abbreviations; 5 - Conventions; 6 -
 25 Source, coded, decoded and output data formats, scanning processes, and neighboring relationships; 7 - Syntax and semantics; 8 - Decoding process and 9 - Parsing process of the HEVC specification.

10. When $sub_pic_cpb_params_present_flag$ is 1 the following relationship shall apply

$$cpb_removal_delay \times t_c = t_{c_sub} \times \sum_{i=0}^{num_decoding_units_minus1} du_cpb_removal_delay[i]$$

The present Annex B lists the seven conditions specified for H.264/AVC in order to achieve bitstream conformance. In order to a bitstream of coded data to be regarded as meeting the requirement of bitstream conformance all of the following conditions should be fulfilled.

- 5 1. For each access unit n , with $n > 0$, associated with a buffering period SEI message, with $\Delta t_{g,90}(n)$ specified by

$$\Delta t_{g,90}(n) = 90000 \times (t_{r,n}(n) - t_{af}(n-1))$$

the value of *initial_cpb_removal_delay*[*SchedSelIdx*] shall be constrained as follows.

If *cbr_flag*[*SchedSelIdx*] is equal to 0,

10
$$\text{initial_cpb_removal_delay}[\text{SchedSelIdx}] \leq \text{Ceil}(\Delta t_{g,90}(n))$$

Otherwise (*cbr_flag*[*SchedSelIdx*] is equal to 1),

$$\text{Floor}(\Delta t_{g,90}(n)) \leq \text{initial_cpb_removal_delay}[\text{SchedSelIdx}] \leq \text{Ceil}(\Delta t_{g,90}(n))$$

- The exact number of bits in the CPB at the removal time of each picture may depend on which buffering period SEI message is selected to initialize the HRD. Encoders must take this into account to
- 15 ensure that all specified constraints must be obeyed regardless of which buffering period SEI message is selected to initialize the HRD, as the HRD may be initialized at any one of the buffering period SEI messages.

2. A CPB overflow is specified as the condition in which the total number of bits in the CPB is larger
- 20 than the CPB size. The CPB shall never overflow.

3. A CPB underflow is specified as the condition in which $t_{r,n}(n)$ is less than $t_{af}(n)$. When *low_delay_hrd_flag* is equal to 0, the CPB shall never underflow.

- 25 4. The nominal removal times of pictures from the CPB (starting from the second picture in decoding order), shall satisfy the constraints on $t_{r,n}(n)$ and $t_r(n)$ expressed in subclauses A.3.1 - Level limits common to the Baseline, Constrained Baseline, Main, and Extended profiles; A.3.2 - Level limits common to the High, High 10, High 4:2:2, High 4:4:4 Predictive, High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, and CAVLC 4:4:4 Intra profiles and A.3.3 - Profile-specific level limits of ISO/IEC 14496-10
- 30 Information technology – Coding of audio-visual objects, Part 10: Advanced Video Coding (denoted H.264/AVC specification herein) for the profile and level specified in the bitstream when decoding a coded video sequence conforming to one or more of the profiles specified in Annex A - Profiles and levels of the H.264/AVC specification using the decoding process specified in clauses 2 - Normative references; 3 - Definitions; 4 - Abbreviations; 5 - Conventions; 6 - Source, coded, decoded and output

data formats, scanning processes, and neighboring relationships; 7 - Syntax and semantics; 8 - Decoding process and 9 - Parsing process of the H.264/AVC specification and they shall satisfy the constraints on $t_{r,n}(n)$ and $t_l(n)$ expressed in subclauses G.10.2.1 - Level limits common to Scalable Baseline, Scalable High, and Scalable High Intra profiles and G.10.2.1 - Profile specific level limits of the H.264/AVC specification for profile and level specified in the bitstream when decoding a coded video sequence conforming to one or more of the profiles specified in Annex G – Scalable video coding of the H.264/AVC specification using the decoding process specified in this Annex G, and they shall satisfy the constraints on $t_{r,n}(n)$ and $t_l(n)$ expressed in subclause H.10.2 – Levels for the profile and level specified in the bitstream when decoding a coded video sequence conforming to one or more of the profiles specified in Annex H – Multiview video coding of the H.264/AVC specification using the decoding process specified in this Annex H.

5. Immediately after any decoded picture is added to the DPB, the fullness of the DPB shall be less than or equal to the DPB size as constrained by Annexes A – Profiles and levels; D – Supplemental enhancement information and E – Video usability information of the H.264/AVC specification and subclauses G.10 – Profiles and levels; G.13 – Supplemental enhancement information; G.14 – Video usability information; H.10 – Profiles and levels; H.13 – MVC SEI messages and H.14 – Video usability information of the H.264/AVC specification for the profile and level specified in the bitstream.

6. All reference pictures shall be present in the DPB when needed for prediction. Each picture shall be present in the DPB at its DPB output time unless it is not stored in the DPB at all, or is removed from the DPB before its output time by one of the processes specified in subclause C.2 - Operation of the decoded picture buffer (DPB) of the H.264/AVC specification.

7. The value of $\Delta t_{o,dpb}(n) = t_{o,dpb}(n_n) - t_{o,dpb}(n)$, which is the difference between the output time of a picture and that of the first picture following it in output order and having *OutputFlag* equal to 1, shall satisfy the constraint expressed in subclause A.3.1 - Level limits common to the Baseline, Constrained Baseline, Main, and Extended profiles of the H.264/AVC specification for the profile, tier and level specified in the bitstream when decoding a coded video sequence conforming to one of more of the profiles specified in Annex A of the H.264/AVC specification using the decoding process specified in clauses 2-9 of the H.264/AVC specification, and it shall satisfy the constraint expressed in subclause G.10.2.1 of the H.264/AVC specification for profile and level specified in the bitstream when decoding a coded video sequence conforming to one or more of the profiles specified in Annex G of the H.264/AVC specification using the decoding process specified in this Annex G, and it shall satisfy the

constraints expressed in subclause H.10.2 of the H.264/AVC specification for the profile and level specified in the bitstream when decoding a coded video sequence conforming to one or more of the profiles specified in Annex H of the H.264/AVC specification using the decoding process specified in this Annex H.

CLAIMS

1. A sub-bitstream extraction method applicable to a bitstream of multiple access units comprising a respective encoded picture (2) of a video sequence (1), said method comprising:
 - identifying (S1) an access unit among said bitstream comprising said multiple access units; and
 - 5 discarding (S2) said access unit and all access units in said bitstream that can depend on said access unit to form a sub-bitstream having no access unit that can depend on said access unit.
2. The method according to claim 1, wherein each picture (2) of said video sequence (1) has a respective layer identifier identifying a respective layer, discarding (S2) said access unit comprises
 - 10 discarding (S2) said access unit and all access units in said bitstream having layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to a decoding order of said bitstream but preceding, according to a defined order of said bitstream, any next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit.
- 15 3. The method according to claim 2, wherein discarding said access unit (S2) comprises:
 - discarding (S2) said access unit from said bitstream; and
 - discarding (S2) all access units in said bitstream having layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to said decoding order
 - 20 but preceding, according to said defined order, a next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit if said bitstream comprises a next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit and following said access unit according to said defined order; or
 - discarding (S2) all access units in said bitstream having layer identifier higher than or equal to
 - 25 said layer identifier of said access unit and following said access unit according to said decoding order if said bitstream does not comprise any next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit and following said access unit according to said defined order.
- 30 4. The method according to claim 2 or 3, wherein discarding (S2) said access unit comprises discarding (S2) said access unit and all access units in said bitstream having layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to a decoding order of said bitstream but preceding, according to said decoding order of said bitstream, any

next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit.

5. The method according to claim 2 or 3, wherein discarding (S2) said access unit comprises
5 discarding (S2) said access unit and all access units in said bitstream having layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to a decoding order of said bitstream but preceding, according to an output order of said bitstream, any next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit.

10

6. The method according to any of the claims 2 to 5, wherein
each picture (2) of said video sequence (1) has a respective temporal identifier identifying a respective temporal layer, and
said layer switching point is a temporal layer switching point selected from a group consisting of
15 a temporal layer access, TLA, picture and an instantaneous decoding refresh, IDR, picture.

7. The method according to any of the claims 1 to 6, wherein discarding (S2) said access unit comprises:
marking (S10) said access unit as to be removed from said bitstream;
20 marking (S10) said all access units in said bitstream that can depend on said access unit as to be removed from said bitstream; and
discarding (S2) all access units of said bitstream marked as to be removed from said bitstream.

8. The method according to any of the claims 1 to 7, wherein discarding (S2) said access unit and
25 said all access units is performed without decoding said access unit.

9. A sub-bitstream extracting device (30) comprising:
an access unit identifier (31) configured to identify an access unit among a bitstream of multiple access units comprising a respective encoded picture (2) of a video sequence (1); and
30 an access unit discarder (32) configured to discard said access unit and all access units in said bitstream that can depend on said access unit to form a sub-bitstream having no access unit that can depend on said access unit.

10. The device according to claim 9, wherein each picture (2) of said video sequence (1) has a respective layer identifier identifying a respective layer, said access unit discarder (32) is configured to discard said access unit and all access units in said bitstream having layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to a decoding
5 order of said bitstream but preceding, according to a defined order of said bitstream, any next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit.

11. The device according to claim 10, wherein said access unit discarder (32) is configured to i)
discard said access unit from said bitstream, and iia) discard all access units in said bitstream having
10 layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to said decoding order but preceding, according to said defined order, a next layer switching point having a layer identifier lower than or equal to said layer identifier of said access unit if
said bitstream comprises a next layer switching point having a layer identifier lower than or equal to
said layer identifier of said access unit and following said access unit according to said defined order,
15 or iib) discard all access units in said bitstream having layer identifier higher than or equal to said layer identifier of said access unit and following said access unit according to said decoding order if said
bitstream does not comprise any next layer switching point having a layer identifier lower than or equal
to said layer identifier of said access unit and following said access unit according to said defined
order.

20

12. The device according to claim 10 or 11, wherein said access unit discarder (32) is configured to
discard said access unit and all access units in said bitstream having layer identifier higher than or
equal to said layer identifier of said access unit and following said access unit according to a decoding
order of said bitstream but preceding, according to said decoding order of said bitstream, any next
25 layer switching point having a layer identifier lower than or equal to said layer identifier of said access
unit.

13. The device according to claim 10 or 11, wherein said access unit discarder (32) is configured to
discard said access unit and all access units in said bitstream having layer identifier higher than or
30 equal to said layer identifier of said access unit and following said access unit according to a decoding
order of said bitstream but preceding, according to an output order of said bitstream, any next layer
switching point having a layer identifier lower than or equal to said layer identifier of said access unit.

14. The device according to any of the claims 9 to 13, further comprising:

a marking unit (33) configured to i) mark said access unit as to be removed from said bitstream, and ii) mark said all access units in said bitstream that can depend on said access unit as to be removed from said bitstream, wherein

said access unit discarder (32) is configured to discard all access units of said bitstream marked
5 as to be removed from said bitstream.

15. The device according to any of the claims 9 to 14, wherein said access unit discarder (32) is configured to discard said access unit and said all access units without decoding said access unit.

10 16. A decoder (40, 50) comprising a sub-bitstream extracting device (30) according to any of the claims 9 to 15.

17. A mobile device (60) comprising a decoder (40, 50) according to claim 16.

15 18. An encoder (70, 80) comprising a sub-bitstream extracting device (30) according to any of the claims 9 to 15.

19. A mobile device (90) comprising an encoder (70, 80) according to claim 18.

20 20. A bitstream extraction method configured to be performed by a network node (80) comprising:
discarding (S20) access units of a bitstream of multiple access units comprising a respective encoded picture (2) of a video sequence (1), each picture (2) of said video sequence (1) has a respective layer identifier identifying a respective layer, discarding (S20) said access units comprises discarding (S20) access units with layer identifier higher than a selected layer identifier X from the
25 bitstream to form a sub-bitstream; and

relaying (S21), to form an updated bitstream, access units with layer identifier higher than X starting from a layer switching point with a layer identifier equal to X or X+1 in said bitstream, wherein discarding (S20) said access units and relaying (S21) said access units are performed without checking said sub-bitstream and said updated bitstream for bitstream conformance.

30

21. The method according to claim 20, wherein

each picture (2) of said video sequence (1) has a respective temporal identifier identifying a respective temporal layer,

said layer switching point is a temporal layer access, TLA, picture, to which it is possible to perform a temporal layer switching, and

relaying (S21) said access units comprises relaying (S21), to form said updated bitstream, said access units with temporal identifier higher than X starting from said TLA picture with a temporal
5 identifier equal to X+1 in said bitstream.

22. The method according to claim 20, wherein

each picture (2) of said video sequence (1) has a respective temporal identifier identifying a respective temporal layer,

10 said layer switching point is a temporal layer access, TLA, picture, from which it is possible to perform a temporal layer switching, and

relaying (S21) said access units comprises relaying (S21), to form said updated bitstream, said access units with temporal identifier higher than X starting from said TLA picture with a temporal identifier equal to X in said bitstream.

15

23. A network node (80) comprising:

an access unit discarder (81) configured to discard access units of a bitstream of multiple access units comprising a respective encoded picture (2) of a video sequence (1), each picture (2) of said video sequence (1) has a respective layer identifier identifying a respective layer, said access unit
20 discarder (81) is configured to discard access units with layer identifier higher than a selected layer identifier X from the bitstream to form a sub-bitstream; and

a relaying unit (82) configured to relay, to form an updated bitstream, access units with layer identifier higher than X starting from a layer switching point with a layer identifier equal to X or X+1 in said bitstream, wherein said access unit discarder (81) is configured to discard said access units and
25 said relaying unit (82) is configured to relay said access units without checking said sub-bitstream and said updated bitstream for bitstream conformance.

24. The network node according to claim 23, wherein

each picture (2) of said video sequence (1) has a respective temporal identifier identifying a
30 respective temporal layer,

said layer switching point is a temporal layer access, TLA, picture, to which it is possible to perform a temporal layer switching, and

said relaying unit (82) is configured to relay, to form said updated bitstream, said access units with temporal identifier higher than X starting from said TLA picture with a temporal identifier equal to X+1 in said bitstream.

5 25. The network node according to claim 23, wherein

each picture (2) of said video sequence (1) has a respective temporal identifier identifying a respective temporal layer,

said layer switching point is a temporal layer access, TLA, picture, from which it is possible to perform a temporal layer switching, and

10 said relaying unit (82) is configured to relay, to form said updated bitstream, said access units with temporal identifier higher than X starting from said TLA picture with a temporal identifier equal to X in said bitstream.

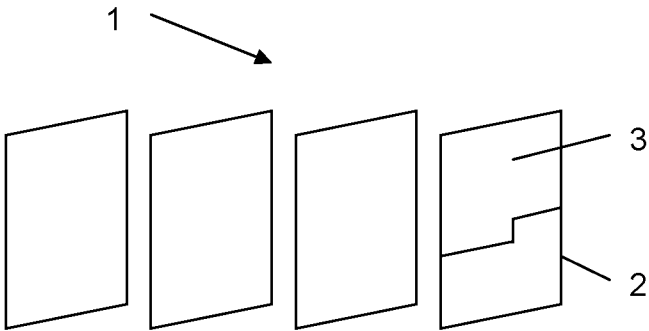


Fig. 1

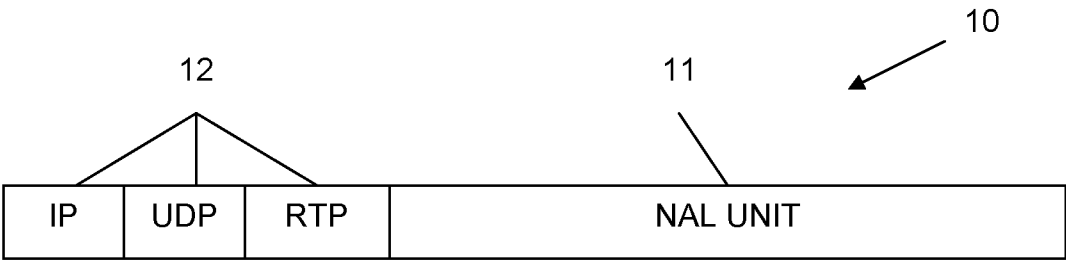


Fig. 2

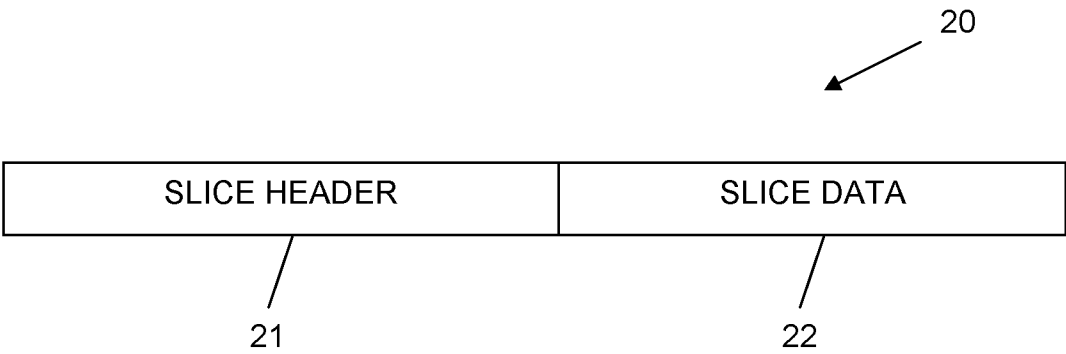


Fig. 3

2/5

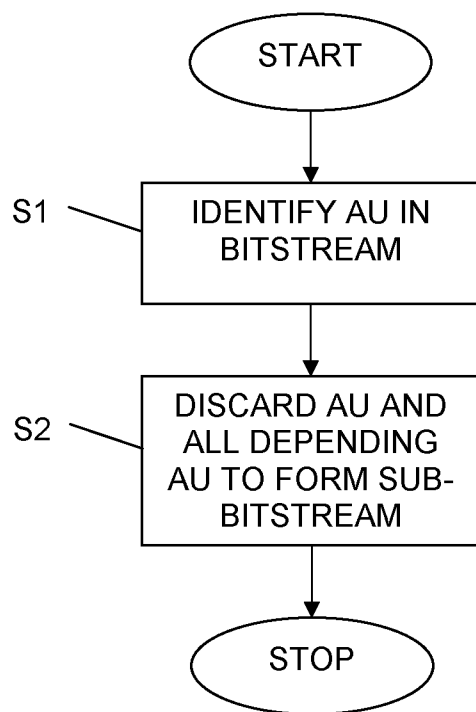


Fig. 4

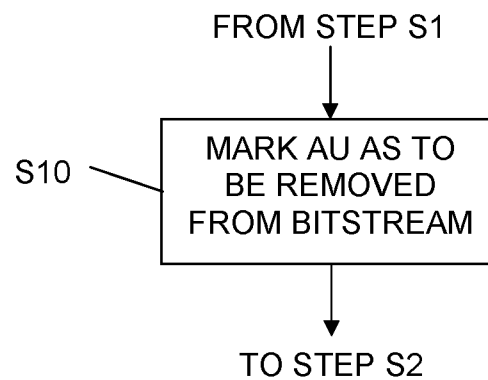


Fig. 5

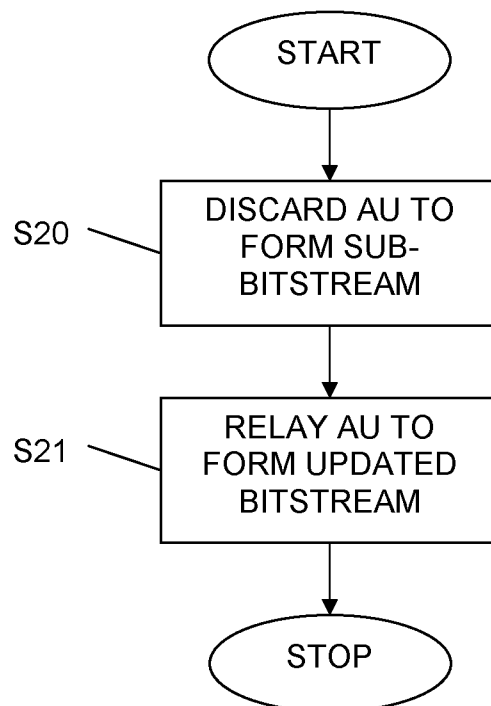


Fig. 13

3/5

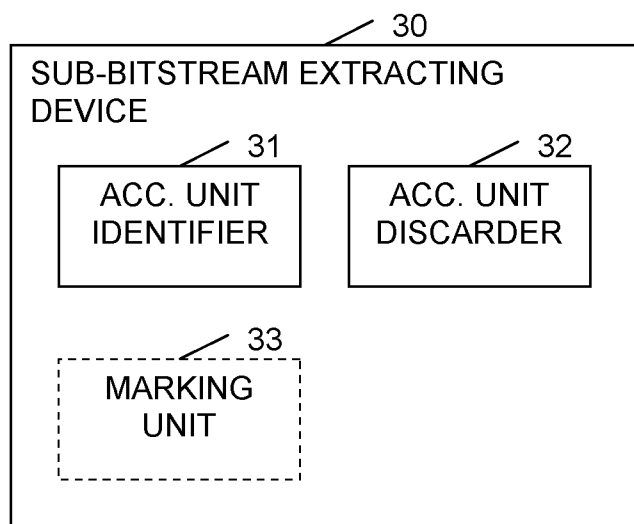


Fig. 6

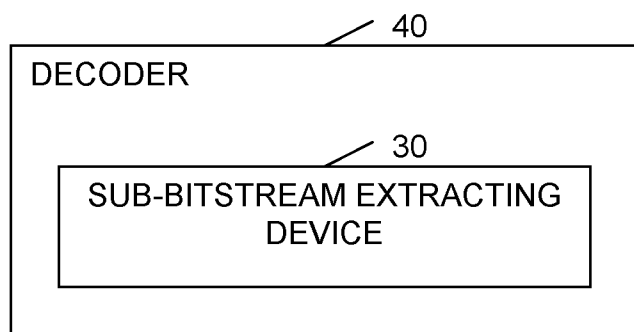


Fig. 7

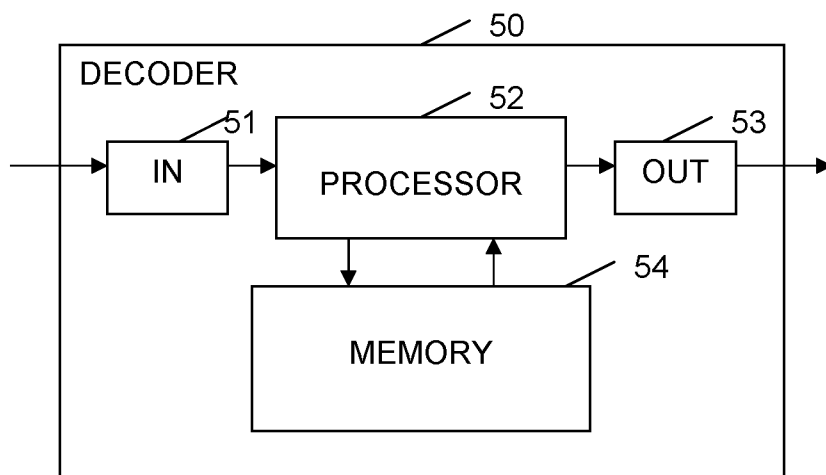


Fig. 8

4/5

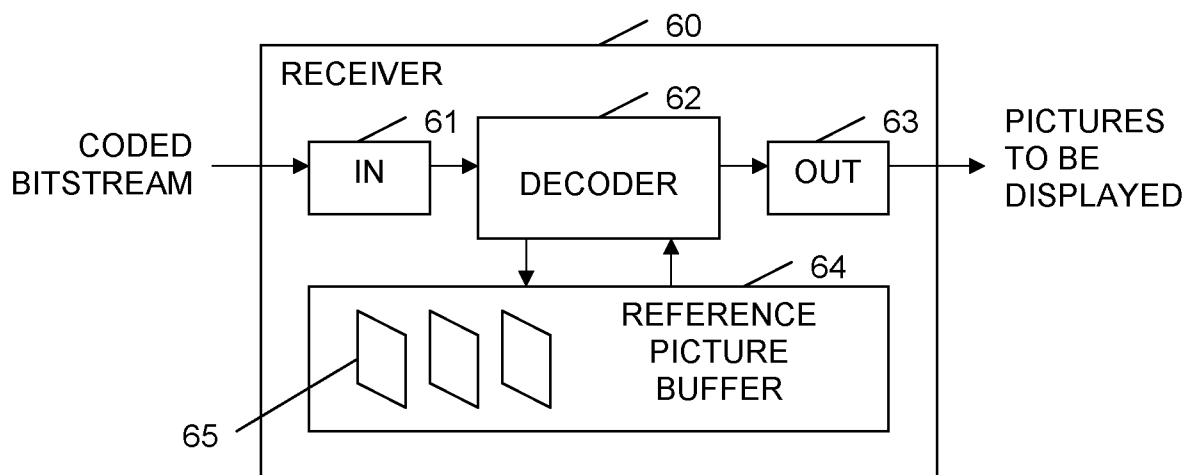


Fig. 9

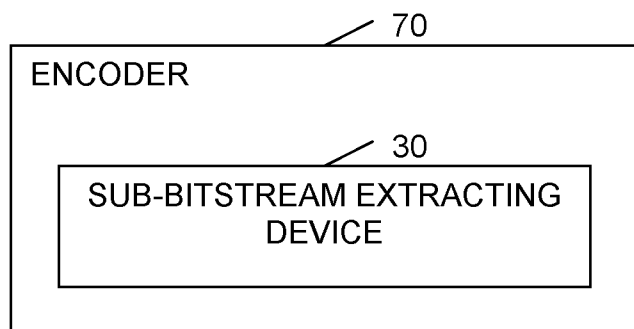


Fig. 10

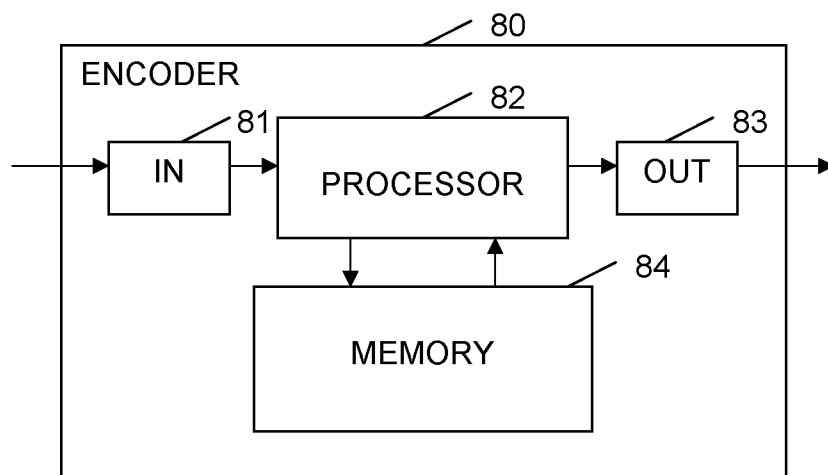


Fig. 11

5/5

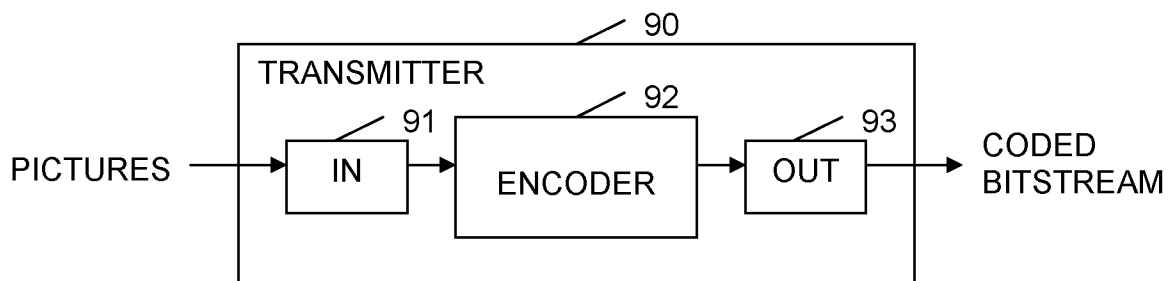


Fig. 12

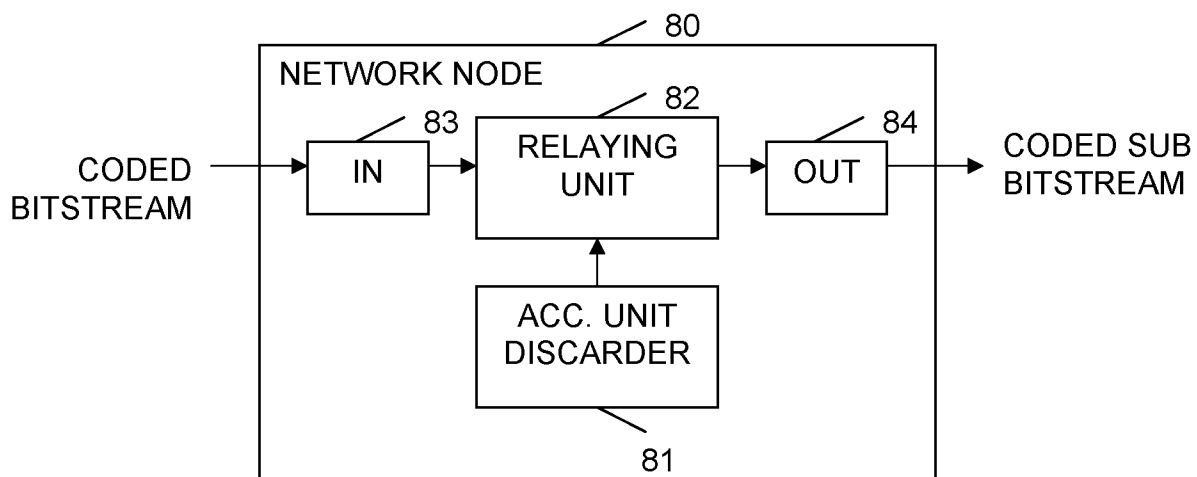


Fig. 14

INTERNATIONAL SEARCH REPORT

International application No
PCT/SE2012/051327

A. CLASSIFICATION OF SUBJECT MATTER

INV. H04N21/2343 H04N21/4402 H04N21/647 H04N21/845 H04N7/26
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EP0-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WENGER INDEPENDENT Y-K WANG HUawei TECHNOLOGIES T SCHIERL FRAUNHOFER HHI A ELEFTHERIADIS VIDYO S: "RTP Payload Format for Scalable Video Coding; rfc6190.txt", RTP PAYLOAD FORMAT FOR SCALABLE VIDEO CODING; RFC6190.TXT, INTERNET ENGINEERING TASK FORCE, IETF; STANDARD, INTERNET SOCIETY (ISOC) 4, RUE DES FALAISES CH- 1205 GENEVA, SWITZERLAND, 7 May 2011 (2011-05-07), pages 1-100, XP015075999, [retrieved on 2011-05-07]	1,2, 7-10, 14-19
Y	pages 9-12 pages 38,39,92 ----- -/--	3-6, 11-13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

6 February 2013

Date of mailing of the international search report

15/02/2013

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Güttlich, Jürgen

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE2012/051327

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 20-25
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No
PCT/SE2012/051327

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SAMUELSSON J ET AL: "Temporal layer access pictures and CRA", 7. JCT-VC MEETING; 98. MPEG MEETING; 21-11-2011 - 30-11-2011; GENEVA; (JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16); URL: HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/, , no. JCTVC-G584, 8 November 2011 (2011-11-08), XP030110568, the whole document -----	3-6, 11-13
X	US 2009/003439 A1 (WANG YE-KUI [FI] ET AL) 1 January 2009 (2009-01-01) paragraphs [0006] - [0034] -----	1,2, 8-10, 15-19
X	US 2010/161823 A1 (JUNG SOON-HEUNG [KR] ET AL) 24 June 2010 (2010-06-24) paragraphs [0006] - [0022], [0033] - [0051] -----	1,2, 8-10,15, 18
X	WO 2008/085013 A1 (UNIV KYUNG HEE UNIV IND COOP [KR]; SUH DOUG YOUNG [KR]; KIM KYU HEON []) 17 July 2008 (2008-07-17) paragraphs [0002] - [0010], [0023], [0024], [0027], [0049], [0123] -----	1,2, 8-10, 16-18
A	BROSS B ET AL: "High Efficiency Video Coding (HEVC) text specification Working Draft 5", 7. JCT-VC MEETING; 98. MPEG MEETING; 21-11-2011 - 30-11-2011; GENEVA; (JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16); URL: HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/, , no. JCTVC-G1103, 19 December 2011 (2011-12-19), XP030111032, pages 24-26 pages 48,56,57 -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/SE2012/051327

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2009003439 A1	01-01-2009	CN 101690229 A EP 2174502 A2 KR 20100030648 A RU 2010102823 A US 2009003439 A1 WO 2009001313 A2	31-03-2010 14-04-2010 18-03-2010 10-08-2011 01-01-2009 31-12-2008
US 2010161823 A1	24-06-2010	KR 20100071688 A US 2010161823 A1	29-06-2010 24-06-2010
WO 2008085013 A1	17-07-2008	CN 101543018 A EP 1994721 A1 KR 20080066624 A US 2009175353 A1 WO 2008085013 A1	23-09-2009 26-11-2008 16-07-2008 09-07-2009 17-07-2008

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 20-25

The independent claims 20 and 23 specify that "access units with layer identifier higher than X" discarded and then "access units with layer identifier higher than X" are relayed. Apparently, no access units for relay purposes are left. The related description page 22 (lines 22-31) states the same. The scope of protection is so unclear that a meaningful search is not possible. This argumentation also relates to the dependent claims 22, 24 and 25 referring to either claim 20 or 23.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.2), should the problems which led to the Article 17(2) declaration be overcome.