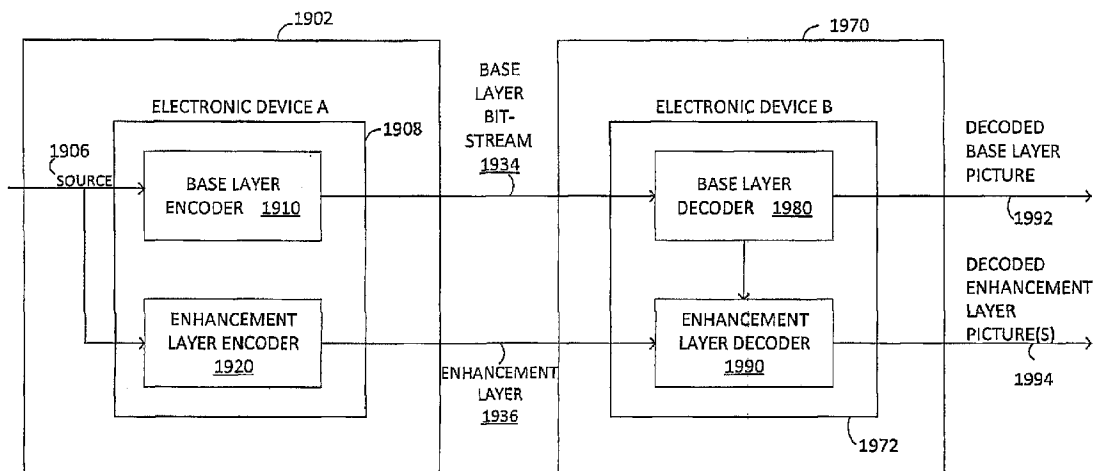




(22) Date de dépôt/Filing Date: 2014/04/04  
(41) Mise à la disp. pub./Open to Public Insp.: 2014/10/16  
(45) Date de délivrance/Issue Date: 2021/10/19  
(62) Demande originale/Original Application: 2 908 888  
(30) Priorités/Priorities: 2013/04/07 (US13/858076);  
2013/07/09 (US61/844272); 2013/07/11 (US61/845309);  
2013/07/19 (US61/856575)

(51) Cl.Int./Int.Cl. *H04N 19/34* (2014.01),  
*H04N 19/174* (2014.01), *H04N 19/423* (2014.01),  
*H04N 19/463* (2014.01), *H04N 19/70* (2014.01)  
(72) Inventeur/Inventor:  
DESHPANDE, SACHIN G., US  
(73) Propriétaire/Owner:  
DOLBY INTERNATIONAL AB, NL  
(74) Agent: SMART & BIGGAR LLP

(54) Titre : SIGNALER UN CHANGEMENT D'ENSEMBLES DE COUCHES DE SORTIE  
(54) Title: SIGNALING CHANGE IN OUTPUT LAYER SETS



(57) Abrégé/Abstract:

A system for decoding a video bitstream includes receiving a bitstream and a plurality of enhancement bitstreams together with receiving a video parameter set and a video parameter set extension. The system also receives an output layer set change message including information indicating a change in at least one output layer set.



### **Abstract**

A system for decoding a video bitstream includes receiving a bitstream and a plurality of enhancement bitstreams together with receiving a video parameter set and a video parameter set extension. The system also receives an output layer set change message including information indicating a change in at least one output layer set.

## **Description**

### **Title of Invention: SIGNALING CHANGE IN OUTPUT LAYER SETS**

This is a divisional application of Canadian National Phase Application No. 2,908,888, filed on 4<sup>th</sup> April, 2014.

### **Technical Field**

[0001] The present disclosure relates generally to electronic devices. More specifically, the present disclosure relates to electronic devices for signaling sub-picture based hypothetical reference decoder parameters, and to systems and methods for hybrid operation of a decoded picture buffer (DPB).

### **Background Art**

[0002] Electronic devices have become smaller and more powerful in order to meet consumer needs and to improve portability and convenience. Consumers have become dependent upon electronic devices and have come to expect increased functionality. Some examples of electronic devices include desktop computers, laptop computers, cellular phones, smart phones, media players, integrated circuits, etc.

[0003] Some electronic devices are used for processing and displaying digital media. For example, portable electronic devices now allow for digital media to be consumed at almost any location where a consumer may be. Furthermore, some electronic devices may provide download or streaming of digital media content for the use and enjoyment of a consumer.

[0004] The increasing popularity of digital media has presented several problems. For example, efficiently representing high-quality digital media for storage, transmittal and rapid playback presents several challenges. As can be observed from this discussion, systems and methods that represent digital media efficiently with improved performance may be beneficial.

[0005] The foregoing and other objectives, features, and advantages of the invention will be more readily understood upon consideration of the following detailed description of the invention, taken in conjunction with the accompanying drawings.

### **Summary of Invention**

[0006] According to an aspect of the present invention, there is provided a method for video decoding, comprising: receiving a bitstream comprising a video parameter set and bits representing a current picture; parsing at least a portion of a slice header of the current picture; performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture; performing a second determination, separate from the first determination, based on one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis; performing the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis; performing the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis; performing a decoding and storing of a current decoded picture in the DPB; performing a third determination, based on the one or more syntax elements in the video parameter set, whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis; marking the current decoded picture in the DPB in accordance with the third determination whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis; and performing an additional picture output from the DPB in accordance with a fourth



determination whether the additional picture output is to be performed on a per-picture basis or per-AU basis.

[0006a] According to another aspect of the present invention, there is provided an electronic device configured for video coding, comprising: a processor; and memory in electronic communication with the processor, wherein instructions stored in the memory are executable to: receive a bitstream comprising a video parameter set and bits representing a current picture, parse at least a portion of a slice header of the current picture, performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture, perform a second determination, separate from the first determination, based on the one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis, perform the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis, perform the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis, perform a decoding and storing of a current decoded picture in the DPB, perform a third determination, based on the one or more syntax elements in the video parameter set, whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis, mark the current decoded picture in the DPB in accordance with the third determination whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis, and perform an additional picture output from the DPB in accordance with a fourth determination whether the additional picture output is to be performed on a per-picture basis or per-AU basis.

[0006b] According to another aspect of the present invention, there is provided a method for video decoding, comprising: receiving a bitstream comprising a video parameter set and bits representing a current picture; parsing at least a portion of a slice header of the current picture; performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture; performing a second determination, separate from the first determination, based on one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis; performing the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis; performing the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis; performing a decoding and storing of a current decoded picture in the DPB.

[0006c] According to another aspect of the present invention, there is provided an electronic device configured for video coding, comprising: a processor; and memory in electronic communication with the processor, wherein instructions stored in the memory are executable to: receive a bitstream comprising a video parameter set and bits representing a current picture, parse at least a portion of a slice header of the current picture, performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture, perform a second determination, separate from the first determination, based on the one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis, perform

the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis, perform the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis, and perform a decoding and storing of a current decoded picture in the DPB.

[0006d] Another aspect provides a method for decoding a video bitstream comprising: (a) receiving a base bitstream representative of a coded video sequence; (b) receiving a plurality of enhancement bitstreams representative of said coded video sequence; (c) receiving a video parameter set containing syntax elements that apply to said base bitstream and said plurality of enhancement bitstreams, wherein said video parameter set contains a syntax element signaling a video parameter set extension; (d) receiving said video parameter set extension containing syntax elements related to at

least one of said enhancement bitstreams; (e) receiving an output layer set change message including information indicating a change in at least one output layer set.

[0007] Another aspect provides a method for decoding a video bitstream comprising: (a) receiving a base bitstream representative of a coded video sequence; (b) receiving a plurality of enhancement bitstreams representative of said coded video sequence; (c) receiving a video parameter set containing syntax elements that apply to said base bitstream and said plurality of enhancement bitstreams, wherein said video parameter set contains a syntax element signaling a video parameter set extension; (d) receiving said video parameter set extension containing syntax elements that includes decoded picture buffer related parameters for a decoded picture buffer for at least one of said enhancement bitstreams.

[0008] Another aspect provides a method for video coding, comprising: beginning to parse a first slice header of a current picture; determining which steps performed by a decoded picture buffer (DPB) will be picture based and which steps will be access unit (AU) based; performing a removal from the DPB; performing a picture output from the DPB; performing a decoding and storing of a current decoded picture in the DPB; marking the current decoded picture in the DPB; and performing an additional picture output from the DPB.

[0009] Another aspect provides an electronic device configured for video coding, comprising: a processor; memory in electronic communication with the processor, wherein instructions stored in the memory are executable to: begin to parse a first slice header of a current picture; determine which steps performed by a decoded picture buffer (DPB) will be picture based and which steps will be access unit (AU) based; perform a removal from the DPB; perform a picture output from the DPB; perform a decoding and storing of a current decoded picture in the DPB; mark the current decoded picture in the DPB; and perform an additional picture output from the DPB.

### **Brief Description of Drawings**

[0010] [fig.1A]Figure 1A is a block diagram illustrating an example of one or more electronic devices in which systems and methods for sending a message and buffering a bitstream may be implemented.

[fig.1B]Figure 1B is another block diagram illustrating an example of one or more electronic devices in which systems and methods for sending a message and buffering a bitstream may be implemented.

[fig.2]Figure 2 is a flow diagram illustrating one configuration of a method for sending a message.

[fig.3]Figure 3 is a flow diagram illustrating one configuration of a method for determining one or more removal delays for decoding units in an access unit.

[fig.4]Figure 4 is a flow diagram illustrating one configuration of a method for buffering a bitstream;

[fig.5]Figure 5 is a flow diagram illustrating one configuration of a method for determining one or more removal delays for decoding units in an access unit.

[fig.6A]Figure 6A is a block diagram illustrating one configuration of an encoder 604 on an electronic device.

[fig.6B]Figure 6B is another block diagram illustrating one configuration of an encoder 604 on an electronic device.

[fig.7A]Figure 7A is a block diagram illustrating one configuration of a decoder on an electronic device.

[fig.7B]Figure 7B is another block diagram illustrating one configuration of a decoder on an electronic device.

[fig.8]Figure 8 illustrates various components that may be utilized in a transmitting electronic device.

[fig.9]Figure 9 is a block diagram illustrating various components that may be utilized in a receiving electronic device.

[fig.10]Figure 10 is a block diagram illustrating one configuration of an electronic device in which systems and methods for sending a message may be implemented.

[fig.11]Figure 11 is a block diagram illustrating one configuration of an electronic device in which systems and methods for buffering a bitstream may be implemented.

[fig.12]Figure 12 is a block diagram illustrating one configuration of a method for operation of a decoded picture buffer.

[fig.13A]Figures 13A illustrates different NAL Unit header syntax.

[fig.13B]Figures 13B illustrates different NAL Unit header syntax.

[fig.13C]Figures 13C illustrates different NAL Unit header syntax.

[fig.14]Figure 14 illustrates a general NAL Unit syntax.

[fig.15]Figure 15 illustrates an existing video parameter set.

[fig.16]Figure 16 illustrates existing scalability types.

[fig.17]Figure 17 illustrates an exemplary video parameter set.

[fig.18]Figure 18 illustrates an exemplary scalability map syntax.

[fig.19]Figure 19 illustrates an exemplary video parameter set.

[fig.20]Figure 20 illustrates an existing video parameter set.

[fig.21]Figure 21 illustrates an existing dimension type, dimension id syntax.

[fig.22]Figure 22 illustrates an exemplary video parameter set.

[fig.23]Figure 23 illustrates an exemplary scalability map syntax.

[fig.24]Figure 24 illustrates an exemplary video parameter set.

[fig.25]Figure 25 illustrates an exemplary video parameter set.

[fig.26]Figure 26 illustrates an exemplary video parameter set.

[fig.27]Figure 27 illustrates an exemplary scalability mask syntax.

[fig.28]Figure 28 illustrates an exemplary video parameter set extension syntax.

[fig.29]Figure 29 illustrates an exemplary video parameter set extension syntax.

[fig.30]Figure 30 illustrates an exemplary video parameter set extension syntax.

[fig.31]Figure 31 illustrates an exemplary video parameter set extension syntax.

[fig.32]Figure 32 illustrates an exemplary video parameter set extension syntax.

[fig.33]Figure 33 illustrates an exemplary video parameter set extension syntax.

[fig.34]Figure 34 illustrates an exemplary video parameter set syntax.

[fig.35]Figure 35 illustrates an exemplary video parameter set extension syntax.

[fig.36]Figure 36 illustrates an exemplary output layer sets change syntax.

[fig.37]Figure 37 illustrates another exemplary output layer sets change syntax.

[fig.38A]Figure 38A illustrates an exemplary video parameter extension syntax.

[fig.38B]Figure 38B illustrates an exemplary video parameter extension syntax.

[fig.39A]Figure 39A illustrates an exemplary op\_dpb\_info\_parameters(j) syntax.

[fig.39B]Figure 39B illustrates an exemplary op\_dpb\_info\_parameters(j) syntax.

[fig.40]Figure 40 illustrates another exemplary video parameter extension syntax.

[fig.41]Figure 41 illustrates another exemplary oop\_dpb\_info\_parameters(j) syntax.

[fig.42]Figure 42 illustrates another exemplary oop\_dpb\_info\_parameters(j) syntax.

[fig.43]Figure 43 illustrates an exemplary num\_dpb\_info\_parameters syntax.

[fig.44]Figure 44 illustrates another exemplary oop\_dpb\_info\_parameters(j) syntax.

[fig.45]Figure 45 illustrates another exemplary num\_dpb\_info\_parameters syntax.

[fig.46]Figure 46 illustrates another exemplary num\_dpb\_info\_parameters syntax.

[fig.47]Figure 47 illustrates another exemplary video parameter extension syntax and layer\_dpb\_info(i).

[fig.48]Figure 48 illustrates an exemplary oop\_dpb\_info\_parameters and layer\_dpb\_info(i) syntax.

[fig.49A]Figure 49A illustrates another exemplary vps\_extension().

[fig.49B]Figure 49B illustrates another exemplary vps\_extension().

[fig.50]Figure 50 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i).

[fig.51]Figure 51 illustrates an exemplary layer\_dpb\_info\_parameters(i).

[fig.52]Figure 52 illustrates another exemplary vps\_extension().

[fig.53]Figure 53 illustrates another exemplary vps\_extension().

[fig.54]Figure 54 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).

[fig.55]Figure 55 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).

[fig.56]Figure 56 illustrates another exemplary vps\_extension().

[fig.57]Figure 57 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).

[fig.58]Figure 58 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).

[fig.59]Figure 59 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).

[fig.60]Figure 60 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).  
 [fig.61]Figure 61 illustrates an exemplary oop\_dpb\_maxbuffering\_parameters(i,k).  
 [fig.62]Figure 62 illustrates an exemplary seq\_parameter\_set\_rbsp().  
 [fig.63]Figure 63 is a block diagram illustrating video coding between multiple electronic devices.  
 [fig.64]Figure 64 is a flow diagram of a method for hybrid decoded picture buffer (DPB) operation.  
 [fig.65]Figure 65 is a flow diagram of another method for hybrid decoded picture buffer (DPB) operation.  
 [fig.66]Figure 66 is a block diagram illustrating one configuration of a decoder;  
 [fig.67A]Figure 67A is a block diagram illustrating the use of both an enhancement layer and a base layer for video coding with separate decoded picture buffers (DPBs) and separate hybrid decoded picture buffer (DPB) operation modules for the base layer and the enhancement layer.  
 [fig.67B]Figure 67B is a block diagram illustrating the use of a shared decoded picture buffer (DPB) and a shared hybrid decoded picture buffer (DPB) operation module for the base layer and the enhancement layer.  
 [fig.68]Figure 68 is a timing diagram illustrating hybrid decoded picture buffer (DPB) operation.  
 [fig.69]Figure 69 is a block diagram illustrating structure and timing for network abstraction layer (NAL) units of layers for coded pictures and access units (AUs) when the second enhancement layer (EL2) has a lower picture rate than the base layer (BL) and the first enhancement layer (EL1).  
 [fig.70]Figure 70 is a block diagram illustrating structure and timing for network abstraction layer (NAL) units of layers for coded pictures and access units (AUs) when the base layer (BL) has a lower picture rate than the first enhancement layer (EL1) and the second enhancement layer (EL2).

## Description of Embodiments

### Example 1

[0011] An electronic device for sending a message is described. The electronic device includes a processor and instructions stored in memory that is in electronic communication with the processor. The electronic device determines, when a Coded Picture Buffer (CPB) supports operation on a sub-picture level, whether to include a common decoding unit CPB removal delay parameter in a picture timing Supplemental Enhancement Information (SEI) message. The electronic device also generates, when the common decoding unit CPB removal delay parameter is to be included in the picture timing SEI message (or some other SEI message or some other parameter set e.g.

picture parameter set or sequence parameter set or video parameter set or adaptation parameter set), the common decoding unit CPB removal delay parameter, wherein the common decoding unit CPB removal delay parameter is applicable to all decoding units in an access unit from the CPB. The electronic device also generates, when the common decoding unit CPB removal delay parameter is not to be included in the picture timing SEI message, a separate decoding unit CPB removal delay parameter for each decoding unit in the access unit. The electronic device also sends the picture timing SEI message with the common decoding unit CPB removal delay parameter or the decoding unit CPB removal delay parameters.

- [0012] The common decoding unit CPB removal delay parameter may specify an amount of sub-picture clock ticks to wait after removal from the CPB of an immediately preceding decoding unit before removing from the CPB a current decoding unit in the access unit associated with the picture timing SEI message.
- [0013] Furthermore, when a decoding unit is a first decoding unit in an access unit, the common decoding unit CPB removal delay parameter may specify an amount of sub-picture clock ticks to wait after removal from the CPB of a last decoding unit in an access unit associated with a most recent buffering period SEI message in a preceding access unit before removing from the CPB the first decoding unit in the access unit associated with the picture timing SEI message.
- [0014] In contrast, when the decoding unit is a non-first decoding unit in an access unit, the common decoding unit CPB removal delay parameter may specify an amount of sub-picture clock ticks to wait after removal from the CPB of a preceding decoding unit in the access unit associated with the picture timing SEI message before removing from the CPB a current decoding unit in the access unit associated with the picture timing SEI message.
- [0015] The decoding unit CPB removal delay parameters may specify an amount of sub-picture clock ticks to wait after removal from the CPB of the last decoding unit before removing from the CPB an i-th decoding unit in the access unit associated with the picture timing SEI message.
- [0016] The electronic device may calculate the decoding unit CPB removal delay parameters according to a remainder of a modulo  $2^{(\text{cpb\_removal\_delay\_length\_minus1} + 1)}$  counter where  $\text{cpb\_removal\_delay\_length\_minus1} + 1$  is a length of a common decoding unit CPB removal delay parameter.
- [0017] The electronic device may also generate, when the CPB supports operation on an access unit level, a picture timing SEI message including a CPB removal delay parameter that specifies how many clock ticks to wait after removal from the CPB of an access unit associated with a most recent buffering period SEI message in a preceding access unit before removing from the CPB the access unit data associated



with the picture timing SEI message.

- [0018] The electronic device may also determine whether the CPB supports operation on a sub-picture level or an access unit level. This may include determining a picture timing flag that indicates whether a Coded Picture Buffer (CPB) provides parameters supporting operation on a sub-picture level based on a value of the picture timing flag. The picture timing flag may be included in the picture timing SEI message.
- [0019] Determining whether to include a common decoding unit CPB removal delay parameter may include setting a common decoding unit CPB removal delay flag to 1 when the common decoding unit CPB removal delay parameter is to be included in the picture timing SEI message. It may also include setting the common decoding unit CPB removal delay flag to 0 when the common decoding unit CPB removal delay parameter is not to be included in the picture timing SEI message. The common decoding unit CPB removal delay flag may be included in the picture timing SEI message.
- [0020] The electronic device may also generate, when the CPB supports operation on a sub-picture level, separate network abstraction layer (NAL) units related parameters that indicate an amount, offset by one, of NAL units for each decoding unit in an access unit. Alternatively, or in addition to, the electronic device may generate a common NAL parameter that indicates an amount, offset by one, of NAL units common to each decoding unit in an access unit.
- [0021] An electronic device for buffering a bitstream is also described. The electronic device includes a processor and instructions stored in memory that is in electronic communication with the processor. The electronic device determines that a CPB signals parameters on a sub-picture level for an access unit. The electronic device also determines, when a received picture timing Supplemental Enhancement Information (SEI) message comprises the common decoding unit Coded Picture Buffer (CPB) removal delay flag, a common decoding unit CPB removal delay parameter applicable to all decoding units in the access unit. The electronic device also determines, when the picture timing SEI message does not comprise the common decoding unit CPB removal delay flag, a separate decoding unit CPB removal delay parameter for each decoding unit in the access unit. The electronic device also removes decoding units from the CPB using the common decoding unit CPB removal delay parameter or the separate decoding unit CPB removal delay parameters. The electronic device also decodes the decoding units in the access unit.
- [0022] In one configuration, the electronic device determines that a picture timing flag is set in the picture timing SEI message. The electronic device may also set a CPB removal delay parameter, `cpb_removal_delay`, according to
- [0023]

[Math.1]

$$\text{cpb\_removal\_delay} = \frac{\left( \sum_{i=0}^{\text{num\_decoding\_units\_minus1}} \text{du\_cpb\_removal\_delay}[i] \right) * t_{c,\text{sub}}}{t_c}$$

[0024] where du\_cpb\_removal\_delay[i] are the decoding unit CPB removal delay parameters,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is a sub-picture clock tick, num\_decoding\_units\_minus1 is an amount of decoding units in the access unit offset by one, and i is an index.

[0025] Alternatively, the electronic device may set a CPB removal delay parameter, cpb\_removal\_delay, and du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] so as to satisfy the equation

[0026] [Math.2]

$$-1 \leq [\text{cpb\_removal\_delay} * t_c - \left( \sum_{i=0}^{\text{num\_decoding\_units\_minus1}} \text{du\_cpb\_removal\_delay}[i] \right) * t_{c,\text{sub}}] \leq 1$$

[0027] where du\_cpb\_removal\_delay[i] are the decoding unit CPB removal delay parameters,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is a sub-picture clock tick, num\_decoding\_units\_minus1 is an amount of decoding units in the access unit offset by one, and i is an index.

[0028] Alternatively, the electronic device may set a CPB removal delay parameter, cpb\_removal\_delay, and du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] according to  $\text{cpb\_removal\_delay} * t_c = \text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}] * t_{c,\text{sub}}$  where du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] is the decoding unit CPB removal delay parameter for the num\_decoding\_units\_minus1'th decoding unit,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is a sub-picture clock tick, num\_decoding\_units\_minus1 is an amount of decoding units in the access unit offset by one.

[0029] In one configuration, the electronic device determines that a picture timing flag is set in the picture timing SEI message. The electronic device may also set CPB removal delay parameters, cpb\_removal\_delay, and du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] so as to satisfy the equation:  $-1 \leq (\text{cpb\_removal\_delay} * t_c - \text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}] * t_{c,\text{sub}}) \leq 1$  where du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] is the decoding unit CPB removal delay parameter for the num\_decoding\_units\_minus1'th decoding unit,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is a sub-picture clock tick, num\_decoding\_units\_minus1 is an amount of decoding units in the access unit offset by one.

[0030] A ClockDiff variable may be defined as  $\text{ClockDiff} = (\text{num\_units\_in\_tick} - (\text{num\_units\_in\_sub\_tick} * (\text{num\_decoding\_units\_minus1} + 1))) /$

time\_scale) where num\_units\_in\_tick is number of time units of a clock operating at the frequency time\_scale Hz that corresponds to one increment of a clock tick counter, num\_units\_in\_sub\_tick is number of time units of a clock operating at the frequency time\_scale Hz that corresponds to one increment of a sub-picture clock tick counter, num\_decoding\_units\_minus1+1 is an amount of decoding units in the access unit, and time\_scale is the number of time units that pass in one second.

- [0031] When a low delay hypothetical reference decoder (HRD) flag (e.g., low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1, the CPB is operating at sub-picture level and ClockDiff is greater than zero, the removal time for decoding unit m,  $t_r(m)$  is determined according to:  $t_r(m) = t_{r,n}(m) + t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub}) + \text{ClockDiff}$  where  $t_{r,n}(m)$  is the nominal removal time of the decoding unit m,  $t_{c\_sub}$  is a sub-picture clock tick, Ceil() is a ceiling function and  $t_{af}(m)$  is final arrival time of decoding unit m.
- [0032] When a low delay hypothetical reference decoder (HRD) flag (e.g., low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1, the CPB is operating at an access unit level and ClockDiff is greater than zero, the removal time for access unit n,  $t_r(n)$  is determined according to:  $t_r(n) = t_{r,n}(n) + t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c) - \text{ClockDiff}$  where  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is a clock tick, Ceil() is a ceiling function and  $t_{af}(n)$  is a final arrival time of access unit n.
- [0033] When a low delay hypothetical reference decoder (HRD) flag (e.g., low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit m of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + \max((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit m,  $t_{c\_sub}$  is sub-picture clock tick, Ceil() is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit m,  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit n.
- [0034] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit n,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + \max((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit n,  $t_{c\_sub}$  is sub-picture clock tick, Ceil() is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit m,  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit n.
- [0035] When a low delay hypothetical reference decoder (HRD) flag (e.g.,

low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit m of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + \min((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit m,  $t_{c\_sub}$  is sub-picture clock tick, Ceil() is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit m,  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit n.

[0036] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit n,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + \min((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit n,  $t_{c\_sub}$  is sub-picture clock tick, Ceil() is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit m,  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit n.

[0037] When a low delay hypothetical reference decoder (HRD) flag (e.g., low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit m of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit m,  $t_{c\_sub}$  is sub-picture clock tick, Ceil() is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit m,  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit n.

[0038] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit n,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit n,  $t_{c\_sub}$  is sub-picture clock tick, Ceil() is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit m,  $t_{r,n}(n)$  is the nominal removal time of the access unit n,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit n.

[0039] When a low delay hypothetical reference decoder (HRD) flag (e.g., low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for a decoding unit m which is not the last decoding unit is set as  $t_r(m) = t_{af}(m)$ , where  $t_{af}(m)$  is a final arrival time of decoding unit m. When a low delay hypothetical reference decoder (HRD) flag (e.g., low\_delay\_hrd\_flag) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for a

decoding unit  $m$  which is the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub}))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick,  $t_{af}(n)$  is a final arrival time of access unit  $n$ , and  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$  in the access unit  $n$ .

- [0040] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for a decoding unit  $m$  which is not the last decoding unit is set as  $t_r(m) = t_{af}(m)$ , where  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ . When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for a decoding unit  $m$  which is the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_c * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick,  $t_{af}(n)$  is a final arrival time of access unit  $n$ , and  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$  in the access unit  $n$ .
- [0041] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for a decoding unit  $m$  is set as  $t_r(m) = t_{af}(m)$  where  $t_{r,n}(m)$  is the nominal removal time of the decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick,  $t_{af}(n)$  is a final arrival time of access unit  $n$ , and  $t_{af}(m)$  is a final arrival time of decoding unit  $m$  in the access unit  $n$ .
- [0042] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{af}(n)$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0043] Additionally in some cases a flag may be sent in part of the bitstream to signal which of the above alternative equations are used for deciding the removal time of the

decoding units and removal time of the access unit. In one case the flag may be called `du_au_cpb_alignment_mode_flag`. If `du_au_cpb_alignment_mode_flag` is 1 then the equations above which align the operation of CPB which operates in sub-picture based mode with the CPB which operates in the access unit mode are used. If

`du_au_cpb_alignment_mode_flag` is 0 then the equations above which do not align the operation of CPB which operates in sub-picture based mode with the CPB which operates in the access unit mode are used.

[0044] In one case the flag `du_au_cpb_alignment_mode_flag` may be signaled in the video usability information (VUI). In another case the flag `du_au_cpb_alignment_mode_flag` may be sent in picture timing SEI message. In yet another case the flag `du_au_cpb_alignment_mode_flag` may be sent in some other normative part of the bitstream. One example of modified syntax and semantics in accordance with the systems and methods disclosed herein is given in Table (0) as follows.

[0045] [Table 0]

<code>pic_timing( payloadSize ) {</code>
<code>  if( CpbDpbDelaysPresentFlag ) {</code>
<code>    cpb_removal_delay</code>
<code>    dpb_output_delay</code>
<code>    if( sub_pic_cpb_params_present_flag ) {</code>
<code>      num_decoding_units_minus1</code>
<code>      du_au_cpb_alignment_mode_flag</code>
<code>      for( i = 0; i &lt;= num_decoding_units_minus1; i++ ) {</code>
<code>        num_nalus_in_du_minus1[i]</code>
<code>        du_cpb_removal_delay[i]</code>
<code>      }</code>
<code>    }</code>
<code>  }</code>
<code>}</code>

Table (0)

[0046] It should be noted that different symbols (names) than those used above for various variables may be used. For example  $t_r(n)$  of access unit  $n$  may be called `CpbRemovalTime(n)`,  $t_d(m)$  of decoding unit  $n$  may be called `CpbRemovalTime(m)`,  $t_{c\_sub}$  may be called `ClockSubTick`,  $t_c$  may be called `ClockTick`,  $t_{ar}(n)$  of access unit  $n$  may be called `FinalArrivalTime(n)` of access unit  $n$ ,  $t_{ar}(m)$  of decoding unit  $m$  may be called `FinalArrivalTime(m)`,  $t_{r,n}(n)$  may be called `NominalRemovalTime(n)` of the access unit  $n$ ,  $t_{r,n}(m)$  may be called `NominalRemovalTime(m)` of the decoding unit  $m$ .

[0047] A method for sending a message by an electronic device is also described. The method includes determining, when a Coded Picture Buffer (CPB) supports operation on a sub-picture level, whether to include a common decoding unit CPB removal delay parameter in a picture timing Supplemental Enhancement Information (SEI) message. The method also includes generating, when the common decoding unit CPB removal delay parameter is to be included in the picture timing SEI message, the common

decoding unit CPB removal delay parameter, wherein the common decoding unit CPB removal delay parameter is applicable to all decoding units in an access unit from the CPB. The method also includes generating, when the common decoding unit CPB removal delay parameter is not to be included in the picture timing SEI message, a separate decoding unit CPB removal delay parameter for each decoding unit in the access unit. The method also includes sending the picture timing SEI message with the common decoding unit CPB removal delay parameter or the decoding unit CPB removal delay parameters.

- [0048] A method for buffering a bitstream by an electronic device is also described. The method includes determining that a CPB signals parameters on a sub-picture level for an access unit. The method also includes determining, when a received picture timing Supplemental Enhancement Information (SEI) message comprises the common decoding unit Coded Picture Buffer (CPB) removal delay flag, a common decoding unit CPB removal delay parameter applicable to all decoding units in the access unit. The method also includes determining, when the picture timing SEI message does not comprise the common decoding unit CPB removal delay flag, a separate decoding unit CPB removal delay parameter for each decoding unit in the access unit. The method also includes removing decoding units from the CPB using the common decoding unit CPB removal delay parameter or the separate decoding unit CPB removal delay parameters. The method also includes decoding the decoding units in the access unit.
- [0049] The systems and methods disclosed herein describe electronic devices for sending a message and buffering a bitstream. For example, the systems and methods disclosed herein describe buffering for bitstreams starting with sub-picture parameters. In some configurations, the systems and methods disclosed herein may describe signaling sub-picture based Hypothetical Reference Decoder (HRD) parameters. For instance, the systems and methods disclosed herein describe modification to a picture timing Supplemental Enhancement Information (SEI) message. The systems and methods disclosed herein (e.g., the HRD modification) may result in more compact signaling of parameters when each sub-picture arrives and is removed from CPB at regular intervals.
- [0050] Furthermore, when the sub-picture level CPB removal delay parameters are present, the Coded Picture Buffer (CPB) may operate at access unit level or sub-picture level. The present systems and methods may also impose a bitstream constraint so that the sub-picture level based CPB operation and the access unit level CPB operation result in the same timing of decoding unit removal. Specifically the timing of removal of last decoding unit in an access unit when operating in sub-picture mode and the timing of removal of access unit when operating in access unit mode will be the same.
- [0051] It should be noted that although the term "hypothetical" is used in reference to an

HRD, the HRD may be physically implemented. For example, "HRD" may be used to describe an implementation of an actual decoder. In some configurations, an HRD may be implemented in order to determine whether a bitstream conforms to High Efficiency Video Coding (HEVC) specifications. For instance, an HRD may be used to determine whether Type I bitstreams and Type II bitstreams conform to HEVC specifications. A Type I bitstream may contain only Video Coding Layer (VCL) Network Access Layer (NAL) units and filler data NAL units. A Type II bitstream may contain additional other NAL units and syntax elements.

[0052] Joint Collaborative Team on Video Coding (JCTVC) document JCTVC-I0333 includes sub-picture based HRD and supports picture timing SEI messages. This functionality has been incorporated into the High Efficiency Video Coding (HEVC) Committee Draft (JCTVC-I1003). B. Bros, W-J. Han, J-R. Ohm, G.J. Sullivan, Wang, and T-. Wiegand, "High efficiency video coding (HEVC) text specification draft 10 (for DFIS & Last Call)", JCTVC- J10003\_v34, Geneva, January 2013, and B. Bros, W-J. Han, J-R. Ohm, G.J. Sullivan, Wang, and T-. Wiegand, "High efficiency video coding (HEVC) text specification draft 10", JCTVC-L1003, Geneva, January 2013. are also referenced herein.

[0053] One example of modified syntax and semantics in accordance with the systems and methods disclosed herein is given in Table (1) as follows.

[0054] [Table 1]

<u>pic_timing</u> (payloadSize) {
if( CpbDpbDelaysPresentFlag ) {
<u>cpb_removal_delay</u>
<u>dpb_output_delay</u>
if( sub_pic_cpb_params_present_flag ) {
<u>num_decoding_units_minus1</u>
<u>common_du_cpb_removal_delay_flag</u>
if( <u>common_du_cpb_removal_delay_flag</u> ) {
<u>common_du_cpb_removal_delay</u>
}
for( i = 0; i <= <u>num_decoding_units_minus1</u> ; i++ ) {
<u>num_nalus_in_du_minus1</u> [i]
if( <u>common_du_cpb_removal_delay_flag</u> )
<u>du_cpb_removal_delay</u> [i]
}
}
}
}

Table (1)

[0055] Examples regarding buffering period SEI message semantics in accordance with the systems and methods disclosed herein are given as follows. In particular, additional detail regarding the semantics of the modified syntax elements are given as follows. When NalHrdBpPresentFlag or VclHrdBpPresentFlag are equal to 1, a buffering



period SEI message can be associated with any access unit in the bitstream, and a buffering period SEI message may be associated with each IDR access unit, with each CRA access unit, and with each access unit associated with a recovery point SEI message. For some applications, the frequent presence of a buffering period SEI message may be desirable. A buffering period is specified as the set of access units between two instances of the buffering period SEI message in decoding order.

- [0056] 'seq\_parameter\_set\_id' specifies the sequence parameter set that contains the sequence HRD attributes. The value of seq\_parameter\_set\_id may be equal to the value of seq\_parameter\_set\_id in the picture parameter set referenced by the primary coded picture associated with the buffering period SEI message. The value of seq\_parameter\_set\_id may be in the range of 0 to 31, inclusive.
- [0057] 'initial\_cpb\_removal\_delay'[ SchedSelIdx ] specifies the delay for the SchedSelIdx-th CPB between the time of arrival in the CPB of the first bit of the coded data associated with the access unit associated with the buffering period SEI message and the time of removal from the CPB of the coded data associated with the same access unit, for the first buffering period after HRD initialization. The syntax element has a length in bits given by initial\_cpb\_removal\_delay\_length\_minus1 + 1. It is in units of a 90 kHz clock. initial\_cpb\_removal\_delay[ SchedSelIdx ] may not be equal to 0 and may not exceed  $90000 * (\text{CpbSize}[\text{SchedSelIdx}] / \text{BitRate}[\text{SchedSelIdx}])$ , the time-equivalent of the CPB size in 90 kHz clock units.
- [0058] 'initial\_cpb\_removal\_delay\_offset'[ SchedSelIdx ] is used for the SchedSelIdx-th CPB in combination with the cpb\_removal\_delay to specify the initial delivery time of coded access units to the CPB. initial\_cpb\_removal\_delay\_offset[ SchedSelIdx ] is in units of a 90 kHz clock. The initial\_cpb\_removal\_delay\_offset[ SchedSelIdx ] syntax element is a fixed length code whose length in bits is given by initial\_cpb\_removal\_delay\_length\_minus1 + 1. This syntax element is not used by decoders and is needed only for the delivery scheduler (HSS) (e.g., as specified in Annex C of JCTVC-I1003).
- [0059] Over the entire coded video sequence, the sum of initial\_cpb\_removal\_delay[ SchedSelIdx ] and initial\_cpb\_removal\_delay\_offset[ SchedSelIdx ] may be constant for each value of SchedSelIdx.
- [0060] 'initial\_du\_cpb\_removal\_delay'[ SchedSelIdx ] specifies the delay for the SchedSelIdx-th CPB between the time of arrival in the CPB of the first bit of the coded data associated with the first decoding unit in the access unit associated with the buffering period SEI message and the time of removal from the CPB of the coded data associated with the same decoding unit, for the first buffering period after HRD initialisation. The syntax element has a length in bits given by initial\_cpb\_removal\_delay\_length\_minus1 + 1. It is in units of a 90 kHz clock.

initial\_du\_cpb\_removal\_delay[ SchedSelIdx ] may not be equal to 0 and may not exceed  $90000 * ( \text{CpbSize[ SchedSelIdx ]} / \text{BitRate[ SchedSelIdx ]} )$ , the time-equivalent of the CPB size in 90 kHz clock units.

- [0061] 'initial\_du\_cpb\_removal\_delay\_offset[ SchedSelIdx ]' is used for the SchedSelIdx-th CPB in combination with the cpb\_removal\_delay to specify the initial delivery time of decoding units to the CPB.
- [0062] initial\_cpb\_removal\_delay\_offset[ SchedSelIdx ] is in units of a 90 kHz clock. The initial\_du\_cpb\_removal\_delay\_offset[ SchedSelIdx ] syntax element is a fixed length code whose length in bits is given by initial\_cpb\_removal\_delay\_length\_minus1 + 1. This syntax element is not used by decoders and is needed only for the delivery scheduler (HSS) (e.g., as specified in Annex C of JCTVC-I1003).
- [0063] Over the entire coded video sequence, the sum of initial\_du\_cpb\_removal\_delay[ SchedSelIdx ] and initial\_du\_cpb\_removal\_delay\_offset[ SchedSelIdx ] may be constant for each value of SchedSelIdx.
- [0064] Examples regarding picture timing SEI message semantics in accordance with the systems and methods disclosed herein are given as follows. In particular, additional detail regarding the semantics of the modified syntax elements are given as follows.
- [0065] The syntax of the picture timing SEI message is dependent on the content of the sequence parameter set that is active for the coded picture associated with the picture timing SEI message. However, unless the picture timing SEI message of an Instantaneous Decoding Refresh (IDR) access unit is preceded by a buffering period SEI message within the same access unit, the activation of the associated sequence parameter set (and, for IDR pictures that are not the first picture in the bitstream, the determination that the coded picture is an IDR picture) does not occur until the decoding of the first coded slice Network Abstraction Layer (NAL) unit of the coded picture. Since the coded slice NAL unit of the coded picture follows the picture timing SEI message in NAL unit order, there may be cases in which it is necessary for a decoder to store the raw byte sequence payload (RBSP) containing the picture timing SEI message until determining the parameters of the sequence parameter that will be active for the coded picture, and then perform the parsing of the picture timing SEI message.
- [0066] The presence of picture timing SEI message in the bitstream is specified as follows. If CpbDpbDelaysPresentFlag is equal to 1, one picture timing SEI message may be present in every access unit of the coded video sequence. Otherwise (CpbDpbDelaysPresentFlag is equal to 0), no picture timing SEI messages may be present in any access unit of the coded video sequence.
- [0067] 'cpb\_removal\_delay' specifies how many clock ticks (see subclause E.2.1 of JCTVC-I1003) to wait after removal from the CPB of the access unit associated with the most

recent buffering period SEI message in a preceding access unit before removing from the buffer the access unit data associated with the picture timing SEI message. This value is also used to calculate an earliest possible time of arrival of access unit data into the CPB for the HSS, as specified in Annex C of JCTVC-I1003. The syntax element is a fixed length code whose length in bits is given by  $\text{cpb\_removal\_delay\_length\_minus1} + 1$ . The  $\text{cpb\_removal\_delay}$  is the remainder of a modulo  $2^{(\text{cpb\_removal\_delay\_length\_minus1} + 1)}$  counter.

- [0068] The value of  $\text{cpb\_removal\_delay\_length\_minus1}$  that determines the length (in bits) of the syntax element  $\text{cpb\_removal\_delay}$  is the value of  $\text{cpb\_removal\_delay\_length\_minus1}$  coded in the sequence parameter set that is active for the primary coded picture associated with the picture timing SEI message, although  $\text{cpb\_removal\_delay}$  specifies a number of clock ticks relative to the removal time of the preceding access unit containing a buffering period SEI message, which may be an access unit of a different coded video sequence.
- [0069] ' $\text{dpb\_output\_delay}$ ' is used to compute the Decoded Picture Buffer (DPB) output time of the picture. It specifies how many clock ticks to wait after removal of the last decoding unit in an access unit from the CPB before the decoded picture is output from the DPB (see subclause C.2 of JCTVC-I1003).
- [0070] With respect to the DPB, a picture is not removed from the DPB at its output time when it is still marked as "used for short-term reference" or "used for long-term reference". Only one  $\text{dpb\_output\_delay}$  is specified for a decoded picture. The length of the syntax element  $\text{dpb\_output\_delay}$  is given in bits by  $\text{dpb\_output\_delay\_length\_minus1} + 1$ . When  $\text{max\_dec\_pic\_buffering}[\text{max\_temporal\_layers\_minus1}]$  is equal to 0,  $\text{dpb\_output\_delay}$  may be equal to 0.
- [0071] The output time derived from the  $\text{dpb\_output\_delay}$  of any picture that is output from an output timing conforming decoder as specified in subclause C.2 of JCTVC-I1003 may precede the output time derived from the  $\text{dpb\_output\_delay}$  of all pictures in any subsequent coded video sequence in decoding order. The picture output order established by the values of this syntax element may be the same order as established by the values of  $\text{PicOrderCnt}()$  as specified by subclause. For pictures that are not output by the "bumping" process of subclause because they precede, in decoding order, an IDR picture with  $\text{no\_output\_of\_prior\_pics\_flag}$  equal to 1 or inferred to be equal to 1, the output times derived from  $\text{dpb\_output\_delay}$  may be increasing with increasing value of  $\text{PicOrderCnt}()$  relative to all pictures within the same coded video sequence.
- [0072] ' $\text{num\_decoding\_units\_minus1}$ ' plus 1 specifies the number of decoding units in the access unit the picture timing SEI message is associated with. The value of  $\text{num\_decoding\_units\_minus1}$  may be in the range of 0 to  $\text{PicWidthInCtbs} * \text{PicHeightInCtbs} - 1$ , inclusive.

- [0073] 'common\_du\_cpb\_removal\_delay\_flag' equal to 1 specifies that the syntax element common\_du\_cpb\_removal\_delay is present.
- [0074] 'common\_du\_cpb\_removal\_delay\_flag' equal to 0 specifies that the syntax element common\_du\_cpb\_removal\_delay is not present.
- [0075] 'common\_du\_cpb\_removal\_delay' specifies information as follows: If a decoding unit is the first decoding unit in the access unit associated with the picture timing SEI message then common\_du\_cpb\_removal\_delay specifies how many sub-picture clock ticks (see subclause E.2.1 of JCTVC-I1003) to wait after removal from the CPB of the last decoding unit in the access unit associated with the most recent buffering period SEI message in a preceding access unit before removing from the CPB the first decoding unit in the access unit associated with the picture timing SEI message.
- [0076] Otherwise common\_du\_cpb\_removal\_delay specifies how many sub-picture clock ticks (see subclause E.2.1 of JCTVC-I1003) to wait after removal from the CPB of the preceding decoding unit in the access unit associated with the picture timing SEI message before removing from the CPB the current decoding unit in the access unit associated with the picture timing SEI message. This value is also used to calculate an earliest possible time of arrival of decoding unit data into the CPB for the HSS, as specified in Annex C. The syntax element is a fixed length code whose length in bits is given by  $\text{cpb\_removal\_delay\_length\_minus1} + 1$ . The common\_du\_cpb\_removal\_delay is the remainder of a modulo  $2^{(\text{cpb\_removal\_delay\_length\_minus1} + 1)}$  counter.
- [0077] An alternate way of specifying 'common\_du\_cpb\_removal\_delay' is as follows.
- [0078] common\_du\_cpb\_removal\_delay specifies how many sub-picture clock ticks (see subclause E.2.1 of JCTVC-I1003) to wait after removal from the CPB of the last decoding unit before removing from the CPB the current decoding unit in the access unit associated with the picture timing SEI message. This value is also used to calculate an earliest possible time of arrival of decoding unit data into the CPB for the HSS, as specified in Annex C. The syntax element is a fixed length code whose length in bits is given by  $\text{cpb\_removal\_delay\_length\_minus1} + 1$ . The common\_du\_cpb\_removal\_delay is the remainder of a modulo  $2^{(\text{cpb\_removal\_delay\_length\_minus1} + 1)}$  counter.
- [0079] The value of  $\text{cpb\_removal\_delay\_length\_minus1}$  that determines the length (in bits) of the syntax element common\_du\_cpb\_removal\_delay is the value of  $\text{cpb\_removal\_delay\_length\_minus1}$  coded in the sequence parameter set that is active for the coded picture associated with the picture timing SEI message, although common\_du\_cpb\_removal\_delay specifies a number of sub-picture clock ticks relative to the removal time of the first decoding unit in the preceding access unit containing a buffering period SEI message, which may be an access unit of a different coded video

sequence.

- [0080] 'num\_nalus\_in\_du\_minus1[ i ]' plus 1 specifies the number of NAL units in the i-th decoding unit of the access unit the picture timing SEI message is associated with. The value of num\_nalus\_in\_du\_minus1[ i ] may be in the range of 0 to PicWidthInCtbs \* PicHeightInCtbs - 1, inclusive.
- [0081] The first decoding unit of the access unit consists of the first num\_nalus\_in\_du\_minus1[ 0 ] + 1 consecutive NAL units in decoding order in the access unit. The i-th (with i greater than 0) decoding unit of the access unit consists of the num\_nalus\_in\_du\_minus1[ i ] + 1 consecutive NAL units immediately following the last NAL unit in the previous decoding unit of the access unit, in decoding order. There may be at least one VCL NAL unit in each decoding unit. All non-VCL NAL units associated with a VCL NAL unit may be included in the same decoding unit.
- [0082] 'du\_cpb\_removal\_delay[ i ]' specifies how many sub-picture clock ticks (see subclause E.2.1 of JCTVC-I1003) to wait after removal from the CPB of the first decoding unit in the access unit associated with the most recent buffering period SEI message in a preceding access unit before removing from the CPB the i-th decoding unit in the access unit associated with the picture timing SEI message. This value is also used to calculate an earliest possible time of arrival of decoding unit data into the CPB for the HSS (e.g., as specified in Annex C of JCTVC-I1003). The syntax element is a fixed length code whose length in bits is given by  $\text{cpb\_removal\_delay\_length\_minus1} + 1$ . The du\_cpb\_removal\_delay[ i ] is the remainder of a modulo  $2^{(\text{cpb\_removal\_delay\_length\_minus1} + 1)}$  counter.
- [0083] The value of cpb\_removal\_delay\_length\_minus1 that determines the length (in bits) of the syntax element du\_cpb\_removal\_delay[ i ] is the value of cpb\_removal\_delay\_length\_minus1 coded in the sequence parameter set that is active for the coded picture associated with the picture timing SEI message, although du\_cpb\_removal\_delay[ i ] specifies a number of sub-picture clock ticks relative to the removal time of the first decoding unit in the preceding access unit containing a buffering period SEI message, which may be an access unit of a different coded video sequence.
- [0084] In one configuration, the timing of decoding unit removal and decoding of decoding units may be implemented as follows.
- [0085] If SubPicCpbFlag is equal to 0, the variable CpbRemovalDelay( m ) is set to the value of cpb\_removal\_delay in the picture timing SEI message associated with the access unit that is decoding unit m, and the variable  $T_c$  is set to  $t_c$ . Otherwise if SubPicCpbFlag is equal to 1 and common\_du\_cpb\_removal\_delay\_flag is 0 the variable CpbRemovalDelay( m ) is set to the value of du\_cpb\_removal\_delay[ i ] for decoding unit m (with m ranging from 0 to num\_decoding\_units\_minus1) in the picture timing

SEI message associated with the access unit that contains decoding unit  $m$ , and the variable  $T_c$  is set to  $t_{c\_sub}$ .

- [0086] In some cases, Otherwise if SubPicCpbFlag is equal to 1 and common\_du\_cpb\_removal\_delay\_flag is 0 the variable CpbRemovalDelay(  $m$  ) is set to the value of  $(m+1)*du\_cpb\_removal\_delay[i]$  for decoding unit  $m$  (with  $m$  ranging from 0 to num\_decoding\_units\_minus1) in the picture timing SEI message associated with the access unit that contains decoding unit  $m$ , and the variable  $T_c$  is set to  $t_{c\_sub}$ .
- [0087] Otherwise if SubPicCpbFlag is equal to 1 and common\_du\_cpb\_removal\_delay\_flag is 1 the variable CpbRemovalDelay(  $m$  ) is set to the value of common\_du\_cpb\_removal\_delay for decoding unit  $m$  in the picture timing SEI message associated with the access unit that contains decoding unit  $m$ , and the variable  $T_c$  is set to  $t_{c\_sub}$ .
- [0088] When a decoding unit  $m$  is the decoding unit with  $n$  equal to 0 (the first decoding unit of the access unit that initializes the HRD), the nominal removal time of the decoding unit from the CPB is specified by  $t_{r,n}(0) = \text{InitCpbRemovalDelay}[\text{SchedSelIdx}] / 90000$ .
- [0089] When a decoding unit  $m$  is the first decoding unit of the first access unit of a buffering period that does not initialize the HRD, the nominal removal time of the decoding unit from the CPB is specified by  $t_{r,n}(m) = t_{r,n}(m_b) + T_c * \text{CpbRemovalDelay}(m)$ , where  $t_{r,n}(m_b)$  is the nominal removal time of the first decoding unit of the previous buffering period.
- [0090] When a decoding unit  $m$  is the first decoding unit of a buffering period,  $m_b$  is set equal to  $m$  at the removal time  $t_{r,n}(m)$  of the decoding unit  $m$ . The nominal removal time  $t_{r,n}(m)$  of a decoding unit  $m$  that is not the first decoding unit of a buffering period is given by  $t_{r,n}(m) = t_{r,n}(m_b) + T_c * \text{CpbRemovalDelay}(m)$ , where  $t_{r,n}(m_b)$  is the nominal removal time of the first decoding unit of the current buffering period.
- [0091] The removal time of decoding unit  $m$  is specified as follows. If low\_delay\_hrd\_flag is equal to 0 or  $t_{r,n}(m) \geq t_{af}(m)$ , the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m)$ . Otherwise (low\_delay\_hrd\_flag is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ), the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m) + T_c * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / T_c)$ . The latter case (low\_delay\_hrd\_flag is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) indicates that the size of decoding unit  $m$ ,  $b(m)$ , is so large that it prevents removal at the nominal removal time.
- [0092] In another case the removal time of decoding unit  $m$  is specified as follows. If low\_delay\_hrd\_flag is equal to 0 or  $t_{r,n}(m) \geq t_{af}(m)$ , the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m)$ . Otherwise (low\_delay\_hrd\_flag is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ), the removal time of decoding unit  $m$  which is not the last decoding unit in the access unit is specified by  $t_r(m) = t_{af}(m)$ , and the removal time

of decoding unit  $m$  which is the last decoding unit in the access unit  $t_r(m) = t_{r,n}(m) + T_c * \text{Ceil}((t_{dr}(m) - t_{r,n}(m)) / T_c)$ . The latter case ( $\text{low\_delay\_hrd\_flag}$  is equal to 1 and  $t_{r,n}(m) < t_{dr}(m)$ ) indicates that the size of decoding unit  $m$ ,  $b(m)$ , is so large that it prevents removal at the nominal removal time.

[0093] In another case the removal time of decoding unit  $m$  is specified as follows. If  $\text{low\_delay\_hrd\_flag}$  is equal to 0 or  $t_{r,n}(m) \geq t_{dr}(m)$ , the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m)$ . Otherwise ( $\text{low\_delay\_hrd\_flag}$  is equal to 1 and  $t_{r,n}(m) < t_{dr}(m)$ ), the removal time of decoding unit  $m$  which is not the last decoding unit in the access unit is specified by  $t_r(m) = t_{dr}(m)$ , and the removal time of decoding unit  $m$  which is the last decoding unit in the access unit  $t_r(m) = t_{r,n}(m) + T_c * \text{Ceil}((t_{dr}(m) - t_{r,n}(m)) / T_c)$ . The latter case ( $\text{low\_delay\_hrd\_flag}$  is equal to 1 and  $t_{r,n}(m) < t_{dr}(m)$ ) indicates that the size of decoding unit  $m$ ,  $b(m)$ , is so large that it prevents removal at the nominal removal time.

[0094] In another case the removal time of decoding unit  $m$  is specified as follows. If  $\text{low\_delay\_hrd\_flag}$  is equal to 0 or  $t_{r,n}(m) \geq t_{dr}(m)$ , the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m)$ . Otherwise ( $\text{low\_delay\_hrd\_flag}$  is equal to 1 and  $t_{r,n}(m) < t_{dr}(m)$ ), the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{dr}(m)$ . The latter case ( $\text{low\_delay\_hrd\_flag}$  is equal to 1 and  $t_{r,n}(m) < t_{dr}(m)$ ) indicates that the size of decoding unit  $m$ ,  $b(m)$ , is so large that it prevents removal at the nominal removal time.

[0095] When  $\text{SubPicCpbFlag}$  is equal to 1, the nominal CPB removal time of access unit  $n$   $t_{r,n}(n)$  is set to the nominal CPB removal time of the last decoding unit in access unit  $n$ , the CPB removal time of access unit  $n$   $t_r(n)$  is set to the CPB removal time of the last decoding unit in access unit  $n$ .

[0096] When  $\text{SubPicCpbFlag}$  is equal to 0, each decoding unit is an access unit, hence the nominal CPB removal time and the CPB removal time of access unit  $n$  are the nominal CPB removal time and the CPB removal time of decoding unit  $n$ .

[0097] At CPB removal time of decoding unit  $m$ , the decoding unit is instantaneously decoded.

[0098] Another example of modified syntax and semantics for a picture timing SEI message in accordance with the systems and methods disclosed herein is given in Table (2) as follows. Modifications in accordance with the systems and methods disclosed herein are denoted in bold.

[0099]

[Table 2]

pic_timing(payloadSize) {
if CpbDpbDelaysPresentFlag) {
cpb_removal_delay
dpb_output_delay
if(sub_pic_cpb_params_present_flag) {
num_decoding_units_minus1
common_du_cpb_removal_delay_flag
if(common_du_cpb_removal_delay_flag) {
common_num_nalus_in_du_minus1
common_du_cpb_removal_delay
}
for(i=0; i<= num_decoding_units_minus1; i++) {
num_nalus_in_du_minus1[i]
if(!common_du_cpb_removal_delay_flag)
du_cpb_removal_delay[i]
}
}
}
}

Table (2)

- [0100] The illustrated example in Table (2) includes a syntax element `common_num_nalus_in_du_minus1`, which may be used to determine how much data should be removed from the CPB when removing a decoding unit. 'common\_num\_nalus\_in\_du\_minus1' plus 1 specifies the number of NAL units in each decoding unit of the access unit the picture timing SEI message is associated with. The value of `common_num_nalus_in_du_minus1` may be in the range of 0 to  $\text{PicWidthInCtbs} * \text{PicHeightInCtbs} - 1$ , inclusive.
- [0101] The first decoding unit of the access unit consists of the first `common_num_nalus_in_du_minus1 + 1` consecutive NAL units in decoding order in the access unit. The  $i$ -th (with  $i$  greater than 0) decoding unit of the access unit consists of the `common_num_nalus_in_du_minus1 + 1` consecutive NAL units immediately following the last NAL unit in the previous decoding unit of the access unit, in decoding order. There may be at least one VCL NAL unit in each decoding unit. All non-VCL NAL units associated with a VCL NAL unit may be included in the same decoding unit.
- [0102] Another example of modified syntax and semantics for a picture timing SEI message in accordance with the systems and methods disclosed herein is given in Table (3) as follows. Modifications in accordance with the systems and methods disclosed herein are denoted in bold.
- [0103]



[Table 3]

pic_timing(payloadSize) {
if( CpbDpbDelaysPresentFlag ) {
cpb_removal_delay
dpb_output_delay
if( sub_pic_cpb_params_present_flag ) {
num_decoding_units_minus1
common_num_nalus_in_du_flag
if( common_num_nalus_in_du_flag ) {
common_num_nalus_in_du_minus1
}
common_du_cpb_removal_delay_flag
if( common_du_cpb_removal_delay_flag ) {
common_du_cpb_removal_delay
}
for( i = 0; i <= num_decoding_units_minus1; i++ ) {
if( !common_num_nalus_in_du_flag )
num_nalus_in_du_minus1[i]
if( !common_du_cpb_removal_delay_flag )
du_cpb_removal_delay[i]
}
}
}
}

Table (3)

- [0104] The illustrated example in Table (3) includes a syntax element 'common\_num\_nalus\_in\_du\_flag' that, when equal to 1, specifies that the syntax element 'common\_num\_nalus\_in\_du\_minus1' is present. 'common\_num\_nalus\_in\_du\_flag' equal to 0 specifies that the syntax element 'common\_num\_nalus\_in\_du\_minus1' is not present.
- [0105] In yet another embodiment flags common\_du\_cpb\_removal\_delay\_flag common\_num\_nalus\_in\_du\_minus1, may not be sent. Instead syntax elements common\_num\_nalus\_in\_du\_minus1 and common\_du\_cpb\_removal\_delay could be sent every time. In this case a value of 0 (or some other) for these syntax elements could be used to indicate that these elements are not signaled.
- [0106] In addition to modifications to the syntax elements and semantics of the picture timing SEI message, the present systems and methods may also implement a bitstream constraint so that sub-picture based CPB operation and access unit level CPB operation result in the same timing of decoding unit removal.
- [0107] When sub\_pic\_cpb\_params\_present\_flag equals to 1 that sub-picture level CPB removal delay parameters are present the CPB may operate at access unit level or sub-picture level. sub\_pic\_cpb\_params\_present\_flag equal to 0 specifies that sub-picture level CPB removal delay parameters are not present and the CPB operates at access unit level. When sub\_pic\_cpb\_params\_present\_flag is not present, its value is inferred to be equal to 0.
- [0108] To support the operation at both access unit level or sub-picture level, the following

bitstream constraints may be used: If sub\_pic\_cpb\_params\_present\_flag is 1 then it is requirement of bitstream conformance that the following constraint is obeyed when signaling the values for cpb\_removal\_delay and du\_cpb\_removal\_delay[i] for all i:

[0109] [Math.3]

$$\text{cpb\_removal\_delay} = \frac{\left( \sum_{i=0}^{\text{num\_decoding\_units\_minus1}} \text{du\_cpb\_removal\_delay}[i] \right) * t_{c,sub}}{t_c}$$

[0110] where du\_cpb\_removal\_delay[i] are the decoding unit CPB removal delay parameters,  $t_c$  is a clock tick,  $t_{c,sub}$  is sub-picture clock tick, num\_decoding\_units\_minus1 is an amount of decoding units in the access unit offset by one, and i is an index. In some embodiments a tolerance parameter could be added to satisfy the above constraint.

[0111] To support the operation at both access unit level or sub-picture level, the bitstream constraints as follows may be used: Let the variable  $T_{du}(k)$  be defined as:

[0112] [Math.4]

$$T_{du}(k) = T_{du}(k-1) + t_{c,sub} * \sum_{i=0}^{\text{num\_decoding\_units\_minus1}_k} (\text{du\_cpb\_removal\_delay\_minus1}_k[i] + 1)$$

[0113] where  $\text{du\_cpb\_removal\_delay\_minus1}_k[i]$  and  $\text{num\_decoding\_units\_minus1}_k$  are parameters for I'th decoding unit of k'th access unit (with  $k=0$  for the access unit that initialized the HRD and  $T_{du}(k)=0$  for  $k<1$ ), and where  $\text{du\_cpb\_removal\_delay\_minus1}_k[i] + 1 = \text{du\_cpb\_removal\_delay\_minus1}_k[i]$  is the decoding unit CPB removal delay parameter for the I'th decoding unit of the k'th access unit, and  $\text{num\_decoding\_units\_minus1}_k$  is the number of decoding units in the k'th access unit,  $t_c$  is a clock tick,  $t_{c,sub}$  is a sub-picture clock tick, and i and k are an indices. Then when the picture timing flag (e.g., sub\_pic\_cpb\_params\_present\_flag) is set to 1, the following constraint shall be true:  $(\text{au\_cpb\_removal\_delay\_minus1} + 1) * t_c = T_{du}(k)$ , where  $(\text{au\_cpb\_removal\_delay\_minus1} + 1) = \text{cpb\_removal\_delay}$ , the CPB removal delay. Thus in this case the CPB removal delay  $(\text{au\_cpb\_removal\_delay\_minus1} + 1)$  is set such that the operation of sub-picture based CPB operation and access unit based CPB operation result in the same timing of access unit removal and last decoding unit of the access unit removal.

[0114] To support the operation at both access unit level or sub-picture level, the following bitstream constraints may be used: If sub\_pic\_cpb\_params\_present\_flag is 1 then it is requirement of bitstream conformance that the following constraint is obeyed when signaling the values for cpb\_removal\_delay and du\_cpb\_removal\_delay[i] for all i:

[0115]

[Math.5]

$$-1 \leq [\text{cpb\_removal\_delay} * t_c - (\sum_{i=0}^{\text{num\_decoding\_units\_minus1}} \text{du\_cpb\_removal\_delay}[i]) * t_{c,\text{sub}}] \leq 1$$

[0116] where  $\text{du\_cpb\_removal\_delay}[i]$  are the decoding unit CPB removal delay parameters,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is sub-picture clock tick,  $\text{num\_decoding\_units\_minus1}$  is an amount of decoding units in the access unit offset by one, and  $i$  is an index.

[0117] To support the operation at both access unit level or sub-picture level, the following bitstream constraints may be used: If  $\text{sub\_pic\_cpb\_params\_present\_flag}$  is 1 then it is requirement of bitstream conformance that the following constraint is obeyed when signaling the values for  $\text{cpb\_removal\_delay}$  and  $\text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}]$ :  $\text{cpb\_removal\_delay} * t_c = \text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}] * t_{c,\text{sub}}$  where  $\text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}]$  is the decoding unit CPB removal delay parameter for the  $\text{num\_decoding\_units\_minus1}$ 'th decoding unit,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is a sub-picture clock tick,  $\text{num\_decoding\_units\_minus1}$  is an amount of decoding units in the access unit offset by one. In some embodiments a tolerance parameter could be added to satisfy the above constraint.

[0118] To support the operation at both access unit level or sub-picture level, the following bitstream constraints may be used: If  $\text{sub\_pic\_cpb\_params\_present\_flag}$  is 1 then it is requirement of bitstream conformance that the following constraint is obeyed when signaling the values for  $\text{cpb\_removal\_delay}$  and  $\text{du\_cpb\_removal\_delay}[i]$  for all  $i$ :  $-1 \leq (\text{cpb\_removal\_delay} * t_c - \text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}] * t_{c,\text{sub}}) \leq 1$  where  $\text{du\_cpb\_removal\_delay}[\text{num\_decoding\_units\_minus1}]$  is the decoding unit CPB removal delay parameter for the  $\text{num\_decoding\_units\_minus1}$ 'th decoding unit,  $t_c$  is a clock tick,  $t_{c,\text{sub}}$  is a sub-picture clock tick,  $\text{num\_decoding\_units\_minus1}$  is an amount of decoding units in the access unit offset by one.

[0119] Additionally, the present systems and methods may modify the timing of decoding unit removal. When sub-picture level CPB removal delay parameters are present, the removal time of decoding unit for "big pictures" (when  $\text{low\_delay\_hrd\_flag}$  is 1 and  $t_{r,n}(m) < t_{af}(m)$ ) may be changed to compensate for difference that can arise due to clock tick counter and sub-picture clock tick counter.

[0120] When  $\text{sub\_pic\_cpb\_params\_present\_flag}$  equals to 1 then sub-picture level CPB removal delay parameters are present and the CPB may operate at access unit level or sub-picture level.  $\text{sub\_pic\_cpb\_params\_present\_flag}$  equal to 0 specifies that sub-picture level CPB removal delay parameters are not present and the CPB operates at access unit level. When  $\text{sub\_pic\_cpb\_params\_present\_flag}$  is not present, its value is inferred to be equal to 0.

- [0121] Specifically, one example of timing of decoding unit removal and decoding of decoding unit implementation is as follows. The variable SubPicCpbPreferredFlag is either specified by external means, or when not specified by external means, set to 0. The variable SubPicCpbFlag is derived as follows: SubPicCpbFlag = SubPicCpbPreferredFlag && sub\_pic\_cpb\_params\_present\_flag. If SubPicCpbFlag is equal to 0, the CPB operates at access unit level and each decoding unit is an access unit. Otherwise the CPB operates at sub-picture level and each decoding unit is a subset of an access unit.
- [0122] If SubPicCpbFlag is equal to 0, the variable CpbRemovalDelay(  $m$  ) is set to the value of cpb\_removal\_delay in the picture timing SEI message associated with the access unit that is decoding unit  $m$ , and the variable  $T_c$  is set to  $t_c$ . Otherwise the variable CpbRemovalDelay(  $m$  ) is set to the value of du\_cpb\_removal\_delay[  $i$  ] for decoding unit  $m$  in the picture timing SEI message associated with the access unit that contains decoding unit  $m$ , and the variable  $T_c$  is set to  $t_{c\_sub}$ .
- [0123] When a decoding unit  $m$  is the decoding unit with  $n$  equal to 0 (the first decoding unit of the access unit that initializes the HRD), the nominal removal time of the decoding unit from the CPB is specified by  $t_{r,n}(0) = \text{InitCpbRemovalDelay}[\text{SchedSelIdx}] / 90000$ .
- [0124] When a decoding unit  $m$  is the first decoding unit of the first access unit of a buffering period that does not initialize the HRD, the nominal removal time of the decoding unit from the CPB is specified by  $t_{r,n}(m) = t_{r,n}(m_b) + T_c * \text{CpbRemovalDelay}(m)$  where  $t_{r,n}(m_b)$  is the nominal removal time of the first decoding unit of the previous buffering period.
- [0125] When a decoding unit  $m$  is the first decoding unit of a buffering period,  $m_b$  is set equal to  $m$  at the removal time  $t_{r,n}(m)$  of the decoding unit  $m$ .
- [0126] The nominal removal time  $t_{r,n}(m)$  of a decoding unit  $m$  that is not the first decoding unit of a buffering period is given by  $t_{r,n}(m) = t_{r,n}(m_b) + T_c * \text{CpbRemovalDelay}(m)$  where  $t_{r,n}(m_b)$  is the nominal removal time of the first decoding unit of the current buffering period.
- [0127] The removal time of decoding unit  $m$  is specified as follows. The variable ClockDiff is defined as  $\text{ClockDiff} = (\text{num\_units\_in\_tick} - (\text{num\_units\_in\_sub\_tick} * (\text{num\_decoding\_units\_minus1} + 1))) / \text{time\_scale}$ . In some case it may be requirement of a bitstream conformance that the parameters num\_units\_in\_tick, num\_units\_in\_sub\_tick, num\_decoding\_units\_minus1 are signaled such that following equation is satisfied.  

$$(\text{num\_units\_in\_tick} - (\text{num\_units\_in\_sub\_tick} * (\text{num\_decoding\_units\_minus1} + 1))) \geq 0$$
- [0128] In some other case it may be requirement of a bitstream conformance that the parameters num\_units\_in\_tick, num\_units\_in\_sub\_tick, num\_decoding\_units\_minus1

may be signaled such that following equation is satisfied.

$$(\text{num\_units\_in\_tick} - (\text{num\_units\_in\_sub\_tick} * (\text{num\_decoding\_units\_minus1} + 1))) \leq 0$$

If `low_delay_hrd_flag` is equal to 0 or  $t_{r,n}(m) \geq t_{af}(m)$ , the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m)$ .

[0129] Otherwise (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ), and when `sub_pic_cpb_params_present_flag` equals to 1 and the CPB is operating at sub-picture level, and if `ClockDiff` is greater than zero the removal time of decoding unit  $m$  when it is the last decoding unit of the access unit  $n$  is specified by  $t_r(m) = t_{r,n}(m) + T_c * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / T_c) + \text{ClockDiff}$ .

[0130] Otherwise (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ), and when `sub_pic_cpb_params_present_flag` equals to 1 and the CPB is operating at access unit level and if `ClockDiff` is less than zero the removal time of access unit  $n$  is specified by  $t_r(m) = t_{r,n}(m) + t_c * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_c) - \text{ClockDiff}$ .

[0131] Otherwise (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ), the removal time of decoding unit  $m$  is specified by  $t_r(m) = t_{r,n}(m) + T_c * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / T_c)$ . The latter case (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) indicates that the size of decoding unit  $m$ ,  $b(m)$ , is so large that it prevents removal at the nominal removal time.

[0132] Otherwise (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) and when a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + \min((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .

[0133] Otherwise (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(n) < t_{af}(n)$ ) and when a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + \min((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .

[0134] Otherwise (`low_delay_hrd_flag` is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) and a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last

decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $Ceil()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .

- [0135] Otherwise ( $low\_delay\_hrd\_flag$  is equal to 1 and  $t_{r,n}(n) < t_{af}(n)$ ) and a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + (t_c * Ceil((t_{af}(n) - t_{r,n}(n)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $Ceil()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0136] Otherwise ( $low\_delay\_hrd\_flag$  is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) and a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the decoding unit which is not the last decoding unit of the access unit is set as  $t_r(m) = t_{af}(m)$ , where  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ . And the removal time of the last decoding unit  $m$  of access unit,  $t_r(m)$  is set according to:  $t_r(m) = t_{r,n}(m) + (t_{c\_sub} * Ceil((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub}))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $Ceil()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$  and  $t_{af}(m)$  is a final arrival time of the last decoding unit  $m$  in the access unit  $n$ .
- [0137] Otherwise ( $low\_delay\_hrd\_flag$  is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) and a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the decoding unit which is not the last decoding unit of the access unit is set as  $t_r(m) = t_{af}(m)$ , where  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ . And the removal time of the last decoding unit  $m$  of access unit,  $t_r(m)$  is set according to:  $t_r(m) = t_{r,n}(m) + (t_c * Ceil((t_{af}(m) - t_{r,n}(m)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $Ceil()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$  and  $t_{af}(m)$  is a final arrival time of the last decoding unit  $m$  in the access unit  $n$ .
- [0138] Otherwise ( $low\_delay\_hrd\_flag$  is equal to 1 and  $t_{r,n}(m) < t_{af}(m)$ ) and a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the decoding unit is set as  $t_r(m) = t_{af}(m)$  where  $t_{r,n}(m)$  is the nominal removal time of the decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $Ceil()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$  and  $t_{af}(m)$  is a final arrival time of the decoding unit  $m$  in the access unit  $n$ .
- [0139] Otherwise ( $low\_delay\_hrd\_flag$  is equal to 1 and  $t_{r,n}(n) < t_{af}(n)$ ) and a picture

timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{at}(n)$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c,sub}$  is sub-picture clock tick,  $Ceil()$  is a ceiling function,  $t_{at}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{at}(n)$  is a final arrival time of access unit  $n$ .

[0140] When SubPicCpbFlag is equal to 1, the nominal CPB removal time of access unit  $n$   $t_{r,n}(n)$  is set to the nominal CPB removal time of the last decoding unit in access unit  $n$ , the CPB removal time of access unit  $n$   $t_r(n)$  is set to the CPB removal time of the last decoding unit in access unit  $n$ .

[0141] When SubPicCpbFlag is equal to 0, each decoding unit is an access unit, hence the nominal CPB removal time and the CPB removal time of access unit  $n$  are the nominal CPB removal time and the CPB removal time of decoding unit  $n$ . At CPB removal time of decoding unit  $m$ , the decoding unit is instantaneously decoded.

[0142] As illustrated by the foregoing, the systems and methods disclosed herein provide syntax and semantics that modify a picture timing SEI message bitstreams carrying sub-picture based parameters. In some configurations, the systems and methods disclosed herein may be applied to HEVC specifications.

[0143] For convenience, several definitions are given as follows, which may be applied to the systems and methods disclosed herein. A random access point may be any point in a stream of data (e.g., bitstream) where decoding of the bitstream does not require access to any point in a bitstream preceding the random access point to decode a current picture and all pictures subsequent to said current picture in output order.

[0144] A buffering period may be specified as a set of access units between two instances of the buffering period SEI message in decoding order. Supplemental Enhancement Information (SEI) may contain information that is not necessary to decode the samples of coded pictures from VCL NAL units. SEI messages may assist in procedures related to decoding, display or other purposes. Conforming decoders may not be required to process this information for output order conformance to HEVC specifications (Annex C of HEVC specifications (JCTVC-I1003) includes specifications for conformance, for example). Some SEI message information may be used to check bitstream conformance and for output timing decoder conformance.

[0145] A buffering period SEI message may be an SEI message related to buffering period. A picture timing SEI message may be an SEI message related to CPB removal timing. These messages may define syntax and semantics which define bitstream arrival timing and coded picture removal timing.

[0146] A Coded Picture Buffer (CPB) may be a first-in first-out buffer containing access units in decoding order specified in a hypothetical reference decoder (HRD). An access

unit may be a set of Network Access Layer (NAL) units that are consecutive in decoding order and contain exactly one coded picture. In addition to the coded slice NAL units of the coded picture, the access unit may also contain other NAL units not containing slices of the coded picture. The decoding of an access unit always results in a decoded picture. A NAL unit may be a syntax structure containing an indication of the type of data to follow and bytes containing that data in the form of a raw byte sequence payload interspersed as necessary with emulation prevention bytes.

[0147] As used herein, the term "common" generally refers to a syntax element or a variable that is applicable to more than one thing. For example, in the context of syntax elements in a picture timing SEI message, the term "common" may mean that the syntax element (e.g., `common_du_cpb_removal_delay`) is applicable to all decoding units in an access unit associated with the picture timing SEI message. Additionally, units of data are described in terms of "n" and "m" generally refer to access units and decoding units, respectively.

[0148] Various configurations are now described with reference to the Figures, where like reference numbers may indicate functionally similar elements. The systems and methods as generally described and illustrated in the Figures herein could be arranged and designed in a wide variety of different configurations. Thus, the following more detailed description of several configurations, as represented in the Figures, is not intended to limit scope, as claimed, but is merely representative of the systems and methods.

[0149] Figure 1A is a block diagram illustrating an example of one or more electronic devices 102 in which systems and methods for sending a message and buffering a bitstream may be implemented. In this example, electronic device A 102a and electronic device B 102b are illustrated. However, it should be noted that one or more of the features and functionality described in relation to electronic device A 102a and electronic device B 102b may be combined into a single electronic device in some configurations.

[0150] Electronic device A 102a includes an encoder 104. The encoder 104 includes a message generation module 108. Each of the elements included within electronic device A 102a (e.g., the encoder 104 and the message generation module 108) may be implemented in hardware, software or a combination of both.

[0151] Electronic device A 102a may obtain one or more input pictures 106. In some configurations, the input picture(s) 106 may be captured on electronic device A 102a using an image sensor, may be retrieved from memory and/or may be received from another electronic device.

[0152] The encoder 104 may encode the input picture(s) 106 to produce encoded data. For example, the encoder 104 may encode a series of input pictures 106 (e.g., video). In



one configuration, the encoder 104 may be a HEVC encoder. The encoded data may be digital data (e.g., part of a bitstream 114). The encoder 104 may generate overhead signaling based on the input signal.

- [0153] The message generation module 108 may generate one or more messages. For example, the message generation module 108 may generate one or more SEI messages or other messages. For a CPB that supports operation on a sub-picture level, the electronic device 102 may send sub-picture parameters, (e.g., CPB removal delay parameter). Specifically, the electronic device 102 (e.g., the encoder 104) may determine whether to include a common decoding unit CPB removal delay parameter in a picture timing SEI message. For example, the electronic device may set a flag (e.g., `common_du_cpb_removal_delay_flag`) to one when the encoder 104 is including a common decoding unit CPB removal delay parameter (e.g., `common_du_cpb_removal_delay`) in the picture timing SEI message. When the common decoding unit CPB removal delay parameter is included, the electronic device may generate the common decoding unit CPB removal delay parameter that is applicable to all decoding units in an access unit. In other words, rather than including a decoding unit CPB removal delay parameter for each decoding unit in an access unit, a common parameter may apply to all decoding units in the access unit with which the picture timing SEI message is associated.
- [0154] In contrast, when the common decoding unit CPB removal delay parameter is not to be included in the picture timing SEI message, the electronic device 102 may generate a separate decoding unit CPB removal delay for each decoding unit in the access unit with which the picture timing SEI message is associated. A message generation module 108 may perform one or more of the procedures described in connection with Figure 2 and Figure 3 below.
- [0155] In some configurations, electronic device A 102a may send the message to electronic device B 102b as part of the bitstream 114. In some configurations electronic device A 102a may send the message to electronic device B 102b by a separate transmission 110. For example, the separate transmission may not be part of the bitstream 114. For instance, a picture timing SEI message or other message may be sent using some out-of-band mechanism. It should be noted that, in some configurations, the other message may include one or more of the features of a picture timing SEI message described above. Furthermore, the other message, in one or more aspects, may be utilized similarly to the SEI message described above.
- [0156] The encoder 104 (and message generation module 108, for example) may produce a bitstream 114. The bitstream 114 may include encoded picture data based on the input picture(s) 106. In some configurations, the bitstream 114 may also include overhead data, such as a picture timing SEI message or other message, slice header(s), PPS(s),

etc. As additional input pictures 106 are encoded, the bitstream 114 may include one or more encoded pictures. For instance, the bitstream 114 may include one or more encoded pictures with corresponding overhead data (e.g., a picture timing SEI message or other message).

- [0157] The bitstream 114 may be provided to a decoder 112. In one example, the bitstream 114 may be transmitted to electronic device B 102b using a wired or wireless link. In some cases, this may be done over a network, such as the Internet or a Local Area Network (LAN). As illustrated in Figure 1A, the decoder 112 may be implemented on electronic device B 102b separately from the encoder 104 on electronic device A 102a. However, it should be noted that the encoder 104 and decoder 112 may be implemented on the same electronic device in some configurations. In an implementation where the encoder 104 and decoder 112 are implemented on the same electronic device, for instance, the bitstream 114 may be provided over a bus to the decoder 112 or stored in memory for retrieval by the decoder 112.
- [0158] The decoder 112 may be implemented in hardware, software or a combination of both. In one configuration, the decoder 112 may be a HEVC decoder. The decoder 112 may receive (e.g., obtain) the bitstream 114. The decoder 112 may generate one or more decoded pictures 118 based on the bitstream 114. The decoded picture(s) 118 may be displayed, played back, stored in memory and/or transmitted to another device, etc.
- [0159] The decoder 112 may include a CPB 120. The CPB 120 may temporarily store encoded pictures. The CPB 120 may use parameters found in a picture timing SEI message to determine when to remove data. When the CPB 120 supports operation on a sub-picture level, individual decoding units may be removed rather than entire access units at one time. The decoder 112 may include a Decoded Picture Buffer (DPB) 122. Each decoded picture is placed in the DPB 122 for being referenced by the decoding process as well as for output and cropping. A decoded picture is removed from the DPB at the later of the DPB output time or the time that it becomes no longer needed for inter-prediction reference.
- [0160] The decoder 112 may receive a message (e.g., picture timing SEI message or other message). The decoder 112 may also determine whether the received message includes a common decoding unit CPB removal delay parameter (e.g., common\_du\_cpb\_removal\_delay). This may include identifying a flag (e.g., common\_du\_cpb\_removal\_delay\_flag) that is set when the common parameter is present in the picture timing SEI message. If the common parameter is present, the decoder 112 may determine the common decoding unit CPB removal delay parameter applicable to all decoding units in the access unit. If the common parameter is not present, the decoder 112 may determine a separate decoding unit CPB removal delay

parameter for each decoding unit in the access unit. The decoder 112 may also remove decoding units from the CPB 120 using either the common decoding unit CPB removal delay parameter or the separate decoding unit CPB removal delay parameters. The CPB 120 may perform one or more of the procedures described in connection with Figure 4 and Figure 5 below.

- [0161] The HRD described above may be one example of the decoder 112 illustrated in Figure 1A. Thus, an electronic device 102 may operate in accordance with the HRD and CPB 120 and DPB 122 described above, in some configurations.
- [0162] It should be noted that one or more of the elements or parts thereof included in the electronic device(s) 102 may be implemented in hardware. For example, one or more of these elements or parts thereof may be implemented as a chip, circuitry or hardware components, etc. It should also be noted that one or more of the functions or methods described herein may be implemented in and/or performed using hardware. For example, one or more of the methods described herein may be implemented in and/or realized using a chipset, an Application-Specific Integrated Circuit (ASIC), a Large-Scale Integrated circuit (LSI) or integrated circuit, etc.
- [0163] Figure 1B is a block diagram illustrating another example of an encoder 1908 and a decoder 1972. In this example, electronic device A 1902 and electronic device B 1970 are illustrated. However, it should be noted that the features and functionality described in relation to electronic device A 1902 and electronic device B 1970 may be combined into a single electronic device in some configurations.
- [0164] Electronic device A 1902 includes the encoder 1908. The encoder 1908 may include a base layer encoder 1910 and an enhancement layer encoder 1920. The video encoder 1908 is suitable for scalable video coding and multi-view video coding, as described later. The encoder 1908 may be implemented in hardware, software or a combination of both. In one configuration, the encoder 1908 may be a high-efficiency video coding (HEVC) coder, including scalable and/or multi-view. Other coders may likewise be used. Electronic device A 1902 may obtain a source 1906. In some configurations, the source 1906 may be captured on electronic device A 1902 using an image sensor, retrieved from memory or received from another electronic device.
- [0165] The encoder 1908 may code the source 1906 to produce a base layer bitstream 1934 and an enhancement layer bitstream 1936. For example, the encoder 1908 may code a series of pictures (e.g., video) in the source 1906. In particular, for scalable video encoding for SNR scalability also known as quality scalability the same source 1906 may be provided to the base layer and the enhancement layer encoder. In particular, for scalable video encoding for spatial scalability a downsampled source may be used for the base layer encoder. In particular, for multi-view encoding a different view source may be used for the base layer encoder and the enhancement layer encoder. The

encoder 1908 may be similar to the encoder 1782 described later in connection with Figure 6B.

- [0166] The bitstreams 1934, 1936 may include coded picture data based on the source 1906. In some configurations, the bitstreams 1934, 1936 may also include overhead data, such as slice header information, PPS information, etc. As additional pictures in the source 1906 are coded, the bitstreams 1934, 1936 may include one or more coded pictures.
- [0167] The bitstreams 1934, 1936 may be provided to the decoder 1972. The decoder 1972 may include a base layer decoder 1980 and an enhancement layer decoder 1990. The video decoder 1972 is suitable for scalable video decoding and multi-view video decoding. In one example, the bitstreams 1934, 1936 may be transmitted to electronic device B 1970 using a wired or wireless link. In some cases, this may be done over a network, such as the Internet or a Local Area Network (LAN). As illustrated in Figure 1B, the decoder 1972 may be implemented on electronic device B 1970 separately from the encoder 1908 on electronic device A 1902. However, it should be noted that the encoder 1908 and decoder 1972 may be implemented on the same electronic device in some configurations. In an implementation where the encoder 1908 and decoder 1972 are implemented on the same electronic device, for instance, the bitstreams 1934, 1936 may be provided over a bus to the decoder 1972 or stored in memory for retrieval by the decoder 1972. The decoder 1972 may provide a decoded base layer 1992 and decoded enhancement layer picture(s) 1994 as output.
- [0168] The decoder 1972 may be implemented in hardware, software or a combination of both. In one configuration, the decoder 1972 may be a high-efficiency video coding (HEVC) decoder, including scalable and/or multi-view. Other decoders may likewise be used. The decoder 1972 may be similar to the decoder 1812 described later in connection with Figure 7B. Also, the base layer encoder and/or the enhancement layer encoder may each include a message generation module, such as that described in relation to Figure 1A. Also, the base layer decoder and/or the enhancement layer decoder may include a coded picture buffer and/or a decoded picture buffer, such as that described in relation to Figure 1A. In addition, the electronic devices of Figure 1B may operate in accordance with the functions of the electronic devices of Figure 1A, as applicable.
- [0169] Figure 2 is a flow diagram illustrating one configuration of a method 200 for sending a message. The method 200 may be performed by an encoder 104 or one of its sub-parts (e.g., a message generation module 108). The encoder 104 may determine 202 a picture timing flag (e.g., sub\_pic\_cpb\_params\_present\_flag) that indicates whether a CPB 120 supports operation on a sub-picture level. For example, when the picture timing flag is set to 1, the CPB 120 may operate on an access unit level or a sub-

picture level. It should be noted that even when the picture timing flag is set to 1, the decision about whether to actually operate at the sub-picture level is left to the decoder 112 itself.

- [0170] The encoder 104 may also determine 204 one or more removal delays for decoding units in an access unit. For example, the encoder 104 may determine a single common decoding unit CPB removal delay parameter (e.g., `common_du_cpb_removal_delay`) that is applicable to all decoding units in the access unit from the CPB 120. Alternatively, the encoder 104 may determine a separate decoding unit CPB removal delay (e.g., `du_cpb_removal_delay[i]`) for each decoding unit in the access unit.
- [0171] The encoder 104 may also determine 206 one or more NAL parameters that indicate an amount, offset by one, of NAL units in each decoding unit in the access point. For example, the encoder 104 may determine a single common NAL parameter (e.g., `common_num_nalus_in_du_minus1`) that is applicable to all decoding units in the access unit from the CPB 120. Alternatively, the encoder 104 may determine a separate decoding unit CPB removal delay (e.g., `num_nalus_in_du_minus1[i]`) for each decoding unit in the access unit.
- [0172] The encoder 104 may also send 208 a picture timing SEI message that includes the picture timing flag, the removal delays and the NAL parameters. The picture timing SEI message may also include other parameters (e.g., `cpb_removal_delay`, `dpb_output_delay`, etc). For example, the electronic device 102 may transmit the message via one or more of wireless transmission, wired transmission, device bus, network, etc. For instance, electronic device A 102a may transmit the message to electronic device B 102b. The message may be part of the bitstream 114, for example. In some configurations, electronic device A 102a may send 208 the message to electronic device B 102b in a separate transmission 110 (that is not part of the bitstream 114).
- [0173] For instance, the message may be sent using some out-of-band mechanism. In some case the information indicated in 204, 206 may be sent in a SEI message different than picture timing SEI message. In yet another case the information indicated in 204, 206 may be sent in a parameter set e.g. video parameter set and/or sequence parameter set and/or picture parameter set and/or adaptation parameter set and/or slice header.
- [0174] Figure 3 is a flow diagram illustrating one configuration of a method 300 for determining one or more removal delays for decoding units in an access unit. In other words, the method 300 illustrated in Figure 3 may further illustrate step 204 in the method 200 illustrated in Figure 2. The method 300 may be performed by an encoder 104. The encoder 104 may determine 302 whether to include a common decoding unit CPB removal delay parameter (e.g., `common_du_cpb_removal_delay`).
- [0175] This may include determining whether a common decoding unit CPB removal delay

flag (e.g., `common_du_cpb_removal_delay_flag`) is set. An encoder 104 may send this common parameter in case the decoding units are removed from the CPB at regular interval. This may be the case, for example, when each decoding unit corresponds to certain number of rows of the picture or has some other regular structure.

- [0176] For example, the common decoding unit CPB removal delay flag may be set to 1 when the common decoding unit CPB removal delay parameter is to be included in the picture timing SEI message and 0 when it is not to be included. If yes (e.g., flag is set to 1), the encoder 104 may determine 304 a common decoding unit CPB removal delay parameter (e.g., `common_du_cpb_removal_delay`) that is applicable to all decoding units in an access unit. If no (e.g., flag is set to 0), the encoder 104 may determine 306 separate decoding unit CPB removal delay parameters (e.g., `du_cpb_removal_delay[i]`) for each decoding unit in an access unit.
- [0177] If a common decoding unit CPB removal delay parameter is present in a picture timing SEI message, it may specify an amount of sub-picture clock ticks to wait after removal from the CPB 120 of an immediately preceding decoding unit before removing from the CPB 120 a current decoding unit in the access unit associated with the picture timing SEI message.
- [0178] For example, when a decoding unit is a first decoding unit in an access unit, the common decoding unit CPB 120 removal delay parameter may specify an amount of sub-picture clock ticks to wait after removal from the CPB 120 of a last decoding unit in an access unit associated with a most recent buffering period SEI message in a preceding access unit before removing from the CPB 120 the first decoding unit in the access unit associated with the picture timing SEI message.
- [0179] When the decoding unit is a non-first decoding unit in an access unit, the common decoding unit CPB removal delay parameter may specify an amount of sub-picture clock ticks to wait after removal from the CPB 120 of a preceding decoding unit in the access unit associated with the picture timing SEI message before removing from the CPB a current decoding unit in the access unit associated with the picture timing SEI message.
- [0180] In contrast, when a common decoding unit CPB removal delay parameter (e.g., `common_du_cpb_removal_delay`) is not sent in a picture timing SEI message, separate decoding unit CPB removal delay parameters (e.g., `du_cpb_removal_delay[i]`) may be included in the picture timing SEI message for each decoding unit in an access unit. The decoding unit CPB removal delay parameters (e.g., `du_cpb_removal_delay[i]`) may specify an amount of sub-picture clock ticks to wait after removal from the CPB 120 of the last decoding unit before removing from the CPB 120 an i-th decoding unit in the access unit associated with the picture timing SEI message. The decoding unit CPB removal delay parameters may be calculated according to a remainder of a

modulo  $2^{(\text{cpb\_removal\_delay\_length\_minus1} + 1)}$  counter where  $\text{cpb\_removal\_delay\_length\_minus1} + 1$  is a length of a common decoding unit CPB removal delay parameter.

- [0181] Figure 4 is a flow diagram illustrating one configuration of a method 400 for buffering a bitstream. The method 400 may be performed by a decoder 112 in an electronic device 102 (e.g., electronic device B 102b), which may receive 402 a message (e.g., a picture timing SEI message or other message). For example, the electronic device 102 may receive 402 the message via one or more of wireless transmission, wired transmission, device bus, network, etc. For instance, electronic device B 102b may receive 402 the message from electronic device A 102a. The message may be part of the bitstream 114, for example. In another example, electronic device B 102b may receive the message from electronic device A 102a in a separate transmission 110 (that is not part of the bitstream 114, for example). For instance, the picture timing SEI message may be received using some out-of-band mechanism. In some configurations, the message may include one or more of a picture timing flag, one or more removal delays for decoding units in an access unit and one or more NAL parameters. Thus, receiving 402 the message may include receiving one or more of a picture timing flag, one or more removal delays for decoding units in an access unit and one or more NAL parameters.
- [0182] The decoder 112 may determine 404 whether a CPB 120 operates on an access unit level or a sub-picture level. For example, a decoder 112 may decide to operate on sub-picture basis if it wants to achieve low latency. Alternatively, the decision may be based on whether the decoder 112 has enough resources to support sub-picture based operation. If the CPB 120 operates on a sub-picture level, the decoder may determine 406 one or more removal delays for decoding units in an access unit. For example, the decoder 112 may determine a single common decoding unit CPB removal delay parameter (e.g.,  $\text{common\_du\_cpb\_removal\_delay}$ ) that is applicable to all decoding units in the access unit. Alternatively, the decoder 112 may determine a separate decoding unit CPB removal delay (e.g.,  $\text{du\_cpb\_removal\_delay}[i]$ ) for each decoding unit in the access unit. In other words, the picture timing SEI message may include a common parameter applicable to all decoding units in an access unit or separate parameters for every decoding unit.
- [0183] The decoder 112 may also remove 408 decoding units based on the removal delays for the decoding units, i.e., using either a common parameter applicable to all decoding units in an access unit or separate parameters for every decoding unit. The decoder 112 may also decode 410 the decoding units.
- [0184] The decoder 112 may use a variable ClockDiff when determining a removal time for determined from various signaled parameters. Specifically, ClockDiff may be determined according to  $\text{ClockDiff} =$

$(\text{num\_units\_in\_tick} - (\text{num\_units\_in\_sub\_tick} * (\text{num\_decoding\_units\_minus1} + 1)) / \text{time\_scale})$  where  $\text{num\_units\_in\_tick}$  is number of time units of a clock operating at the frequency  $\text{time\_scale}$  Hz that corresponds to one increment of a clock tick counter,  $\text{num\_units\_in\_sub\_tick}$  is number of time units of a clock operating at the frequency  $\text{time\_scale}$  Hz that corresponds to one increment of a sub-picture clock tick counter,  $\text{num\_decoding\_units\_minus1} + 1$  is an amount of decoding units in the access unit, and  $\text{time\_scale}$  is the number of time units that pass in one second.

- [0185] When a low delay hypothetical reference decoder (HRD) flag (e.g.,  $\text{low\_delay\_hrd\_flag}$ ) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1, the CPB is operating at sub-picture level and  $\text{ClockDiff}$  is greater than zero, the removal time for decoding unit  $m$ ,  $t_r(m)$  is determined according to:  $t_r(m) = t_{r,n}(m) + t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub}) + \text{ClockDiff}$  where  $t_{r,n}(m)$  is the nominal removal time of the decoding unit  $m$ ,  $t_{c\_sub}$  is a sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function and  $t_{af}(m)$  is final arrival time of decoding unit  $m$ .
- [0186] When a low delay hypothetical reference decoder (HRD) flag (e.g.,  $\text{low\_delay\_hrd\_flag}$ ) is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1, the CPB is operating at an access unit level and  $\text{ClockDiff}$  is greater than zero, the removal time for access unit  $n$ ,  $t_r(n)$  is determined according to:  $t_r(n) = t_{r,n}(n) + t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c) - \text{ClockDiff}$  where  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is a clock tick,  $\text{Ceil}()$  is a ceiling function and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0187] When a low delay hypothetical reference decoder (HRD) flag (e.g.,  $\text{low\_delay\_hrd\_flag}$ ) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + \max((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0188] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + \max((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .



- [0189] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + \min((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0190] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + \min((t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub})), (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c)))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0191] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0192] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{r,n}(n) + (t_c * \text{Ceil}((t_{af}(n) - t_{r,n}(n)) / t_c))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .
- [0193] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the decoding unit which is not the last decoding unit of the access unit is set as  $t_r(m) = t_{af}(m)$ , where  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ . And the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_{c\_sub} * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub}))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is

sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick,  $t_{af}(n)$  is a final arrival time of access unit  $n$  and  $t_{af}(m)$  is a final arrival time of the last decoding unit in the access unit  $n$ .

[0194] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the decoding unit which is not the last decoding unit of the access unit is set as  $t_r(m) = t_{af}(m)$ , where  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ . And the removal time for the last decoding unit  $m$  of access unit,  $t_r(m)$  according to:  $t_r(m) = t_{r,n}(m) + (t_c * \text{Ceil}((t_{af}(m) - t_{r,n}(m)) / t_{c\_sub}))$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick,  $t_{af}(n)$  is a final arrival time of access unit  $n$  and  $t_{af}(m)$  is a final arrival time of the last decoding unit in the access unit  $n$ .

[0195] When a low delay hypothetical reference decoder (HRD) flag (e.g., `low_delay_hrd_flag`) is set to 1,  $t_{r,n}(m) < t_{af}(m)$ , a picture timing flag is set to 1 and the CPB is operating at sub-picture level, the removal time for the decoding unit is set as  $t_r(m) = t_{af}(m)$  where  $t_{r,n}(m)$  is the nominal removal time of the decoding unit  $m$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick,  $t_{af}(n)$  is a final arrival time of access unit  $n$  and  $t_{af}(m)$  is a final arrival time of the decoding unit in the access unit  $n$ .

[0196] When a low delay hypothetical reference decoder (HRD) flag is set to 1,  $t_{r,n}(n) < t_{af}(n)$ , a picture timing flag is set to 1 and the CPB is operating at access unit level, the removal time for access unit  $n$ ,  $t_r(n)$  according to:  $t_r(n) = t_{af}(n)$  where  $t_{r,n}(m)$  is the nominal removal time of the last decoding unit  $n$ ,  $t_{c\_sub}$  is sub-picture clock tick,  $\text{Ceil}()$  is a ceiling function,  $t_{af}(m)$  is a final arrival time of last decoding unit  $m$ ,  $t_{r,n}(n)$  is the nominal removal time of the access unit  $n$ ,  $t_c$  is clock tick and  $t_{af}(n)$  is a final arrival time of access unit  $n$ .

[0197] If the CPB operates on an access unit level, the decoder 112 may determine 412 a CPB removal delay parameter. This may be included in the received picture timing SEI message (e.g., `cpb_removal_delay`). The decoder 112 may also remove 414 an access unit based on the CPB removal delay parameter and decode 416 the access unit. In other words, the decoder 112 may decode whole access units at a time, rather than decoding units within the access unit.

[0198] Figure 5 is a flow diagram illustrating one configuration of a method 500 for determining one or more removal delays for decoding units in an access unit. In other

words, the method 500 illustrated in Figure 5 may further illustrate step 406 in the method 400 illustrated in Figure 4. The method 500 may be performed by a decoder 112. The decoder 112 may determine 502 whether a received picture timing SEI message includes a common decoding unit CPB removal delay parameter. This may include determining whether a common decoding unit CPB removal delay flag (e.g., common\_du\_cpb\_removal\_delay\_flag) is set. If yes, the decoder 112 may determine 504 a common decoding unit CPB removal delay parameter (e.g., common\_du\_cpb\_removal\_delay) that is applicable to all decoding units in an access unit. If no, the decoder 112 may determine 506 separate decoding unit CPB removal delay parameters (e.g., du\_cpb\_removal\_delay[i]) for each decoding unit in an access unit.

[0199] In addition to modifying the picture timing SEI message semantics, the present systems and methods may also impose a bitstream constraint so that the operation of sub-picture based CPB operation and access unit based CPB operation result in the same timing of decoding unit removal. Specifically, when the picture timing flag (e.g., sub\_pic\_cpb\_params\_present\_flag) is set to 1, the CPB removal delay may be set according to

[0200] [Math.6]

$$\text{cpb\_removal\_delay} = \frac{\left( \sum_{i=0}^{\text{num\_decoding\_units\_minus1}} \text{du\_cpb\_removal\_delay}[i] \right) * t_{c,sub}}{t_c}$$

[0201] where du\_cpb\_removal\_delay[i] are the decoding unit CPB removal delay parameters,  $t_c$  is a clock tick,  $t_{c,sub}$  is a sub-picture clock tick, num\_decoding\_units\_minus1 is an amount of decoding units in the access unit offset by one, and i is an index.

[0202] Alternatively, the CPB removal delay may be set as described next: Let the variable  $T_{du}(k)$  be defined as:

[0203] [Math.7]

$$T_{du}(k) = T_{du}(k-1) + t_{c,sub} * \sum_{i=0}^{\text{num\_decoding\_units\_minus1}_k} (\text{du\_cpb\_removal\_delay\_minus1}_k[i] + 1)$$

[0204] where du\_cpb\_removal\_delay\_minus1<sub>k</sub>[i] and num\_decoding\_units\_minus1<sub>k</sub> are parameters for I'th decoding unit of k'th access unit (with k=0 for the access unit that initialized the HRD and  $T_{du}(k)=0$  for  $k<1$ ), and where du\_cpb\_removal\_delay\_minus1<sub>k</sub>[i]+1=du\_cpb\_removal\_delay\_minus1<sub>k</sub>[i] is the decoding unit CPB removal delay parameter for the I'th decoding unit of the k'th access unit, and num\_decoding\_units\_minus1<sub>k</sub> is the number of decoding units in the k'th access unit,  $t_c$  is a clock tick,  $t_{c,sub}$  is a sub-picture clock tick, and i and k are an indices. Then when

the picture timing flag (e.g., sub\_pic\_cpb\_params\_present\_flag) is set to 1, the following condition shall be true:

- [0205]  $(au\_cpb\_removal\_delay\_minus1 + 1) * t_c = T_{du}(k)$ , where  
 $(au\_cpb\_removal\_delay\_minus1 + 1) = cpb\_removal\_delay$ , the CPB removal delay.  
 Thus in this case the CPB removal delay  $(au\_cpb\_removal\_delay\_minus1 + 1)$  is set such that the operation of sub-picture based CPB operation and access unit based CPB operation result in the same timing of access unit removal and last decoding unit of the access unit removal.

- [0206] Alternatively, the CPB removal delay may be set according to

- [0207] [Math.8]

$$-1 \leq cpb\_removal\_delay * t_c - \left( \sum_{i=0}^{num\_decoding\_units\_minus1} du\_cpb\_removal\_delay[i] * t_{c,sub} \right) \leq 1$$

- [0208] where  $du\_cpb\_removal\_delay[i]$  are the decoding unit CPB removal delay parameters,  $t_c$  is a clock tick,  $t_{c,sub}$  is a sub-picture clock tick,  $num\_decoding\_units\_minus1$  is an amount of decoding units in the access unit offset by one, and  $i$  is an index.

- [0209] Alternatively,  $cpb\_removal\_delay$  and  $du\_cpb\_removal\_delay[num\_decoding\_units\_minus1]$  may be set according to:  
 $cpb\_removal\_delay * t_c = du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] * t_{c,sub}$   
 where  $du\_cpb\_removal\_delay[num\_decoding\_units\_minus1]$  is the decoding unit CPB removal delay parameter for the  $num\_decoding\_units\_minus1$ 'th decoding unit,  $t_c$  is a clock tick,  $t_{c,sub}$  is a sub-picture clock tick,  $num\_decoding\_units\_minus1$  is an amount of decoding units in the access unit offset by one.

- [0210] In addition to modifying the picture timing SEI message semantics, the present systems and methods may also impose a bitstream constraint so that the operation of sub-picture based CPB operation and access unit based CPB operation result in the same timing of decoding unit removal. Specifically, when the picture timing flag (e.g., sub\_pic\_cpb\_params\_present\_flag) is set to 1, the values for  $cpb\_removal\_delay$  and  $du\_cpb\_removal\_delay[num\_decoding\_units\_minus1]$  may be set so as to satisfy: -  
 $1 \leq (cpb\_removal\_delay * t_c - du\_cpb\_removal\_delay[num\_decoding\_units\_minus1] * t_{c,sub}) \leq 1$  where  $du\_cpb\_removal\_delay[num\_decoding\_units\_minus1]$  is the decoding unit CPB removal delay parameter for the  $num\_decoding\_units\_minus1$ 'th decoding unit,  $t_c$  is a clock tick,  $t_{c,sub}$  is a sub-picture clock tick,  $num\_decoding\_units\_minus1$  is an amount of decoding units in the access unit offset by one.

- [0211] Figure 6A is a block diagram illustrating one configuration of an encoder 604 on an electronic device 602. It should be noted that one or more of the elements illustrated as included within the electronic device 602 may be implemented in hardware, software

or a combination of both. For example, the electronic device 602 includes an encoder 604, which may be implemented in hardware, software or a combination of both. For instance, the encoder 604 may be implemented as a circuit, integrated circuit, application-specific integrated circuit (ASIC), processor in electronic communication with memory with executable instructions, firmware, field-programmable gate array (FPGA), etc., or a combination thereof. In some configurations, the encoder 604 may be a HEVC coder.

- [0212] The electronic device 602 may include a source 622. The source 622 may provide picture or image data (e.g., video) as one or more input pictures 606 to the encoder 604. Examples of the source 622 may include image sensors, memory, communication interfaces, network interfaces, wireless receivers, ports, etc.
- [0213] One or more input pictures 606 may be provided to an intra-frame prediction module and reconstruction buffer 624. An input picture 606 may also be provided to a motion estimation and motion compensation module 646 and to a subtraction module 628.
- [0214] The intra-frame prediction module and reconstruction buffer 624 may generate intra mode information 640 and an intra-signal 626 based on one or more input pictures 606 and reconstructed data 660. The motion estimation and motion compensation module 646 may generate inter mode information 648 and an inter signal 644 based on one or more input pictures 606 and a reference picture 678 from decoded picture buffer 676. In some configurations, the decoded picture buffer 676 may include data from one or more reference pictures in the decoded picture buffer 676.
- [0215] The encoder 604 may select between the intra signal 626 and the inter signal 644 in accordance with a mode. The intra signal 626 may be used in order to exploit spatial characteristics within a picture in an intra-coding mode. The inter signal 644 may be used in order to exploit temporal characteristics between pictures in an inter coding mode. While in the intra coding mode, the intra signal 626 may be provided to the subtraction module 628 and the intra mode information 640 may be provided to an entropy coding module 642. While in the inter coding mode, the inter signal 644 may be provided to the subtraction module 628 and the inter mode information 648 may be provided to the entropy coding module 642.
- [0216] Either the intra signal 626 or the inter signal 644 (depending on the mode) is subtracted from an input picture 606 at the subtraction module 628 in order to produce a prediction residual 630. The prediction residual 630 is provided to a transformation module 632. The transformation module 632 may compress the prediction residual 630 to produce a transformed signal 634 that is provided to a quantization module 636. The quantization module 636 quantizes the transformed signal 634 to produce transformed and quantized coefficients (TQCs) 638.
- [0217] The TQCs 638 are provided to an entropy coding module 642 and an inverse quan-

tization module 650. The inverse quantization module 650 performs inverse quantization on the TQCs 638 to produce an inverse quantized signal 652 that is provided to an inverse transformation module 654. The inverse transformation module 654 decompresses the inverse quantized signal 652 to produce a decompressed signal 656 that is provided to a reconstruction module 658.

- [0218] The reconstruction module 658 may produce reconstructed data 660 based on the decompressed signal 656. For example, the reconstruction module 658 may reconstruct (modified) pictures. The reconstructed data 660 may be provided to a deblocking filter 662 and to the intra prediction module and reconstruction buffer 624. The deblocking filter 662 may produce a filtered signal 664 based on the reconstructed data 660.
- [0219] The filtered signal 664 may be provided to a sample adaptive offset (SAO) module 666. The SAO module 666 may produce SAO information 668 that is provided to the entropy coding module 642 and an SAO signal 670 that is provided to an adaptive loop filter (ALF) 672. The ALF 672 produces an ALF signal 674 that is provided to the decoded picture buffer 676. The ALF signal 674 may include data from one or more pictures that may be used as reference pictures.
- [0220] The entropy coding module 642 may code the TQCs 638 to produce bitstream A 614a (e.g., encoded picture data). For example, the entropy coding module 642 may code the TQCs 638 using Context-Adaptive Variable Length Coding (CAVLC) or Context-Adaptive Binary Arithmetic Coding (CABAC). In particular, the entropy coding module 642 may code the TQCs 638 based on one or more of intra mode information 640, inter mode information 648 and SAO information 668. Bitstream A 614a (e.g., encoded picture data) may be provided to a message generation module 608. The message generation module 608 may be configured similarly to the message generation module 108 described in connection with Figure 1. Additionally or alternatively, the message generation module 608 may perform one or more of the procedures described in connection with Figure 2 and Figure 3.
- [0221] For example, the message generation module 608 may generate a message (e.g., picture timing SEI message or other message) including sub-picture parameters. The sub-picture parameters may include one or more removal delays for decoding units (e.g., `common_du_cpb_removal_delay` or `du_cpb_removal_delay[i]`) and one or more NAL parameters (e.g., `common_num_nalus_in_du_minus1` or `num_nalus_in_du_minus1[i]`). In some configurations, the message may be inserted into bitstream A 614a to produce bitstream B 614b. Thus, the message may be generated after the entire bitstream A 614a is generated (e.g., after most of bitstream B 614b is generated), for example. In other configurations, the message may not be inserted into bitstream A 614a (in which case bitstream B 614b may be the same as bitstream A 614a), but may be provided in a separate transmission 610.

- [0222] In some configurations, the electronic device 602 sends the bitstream 614 to another electronic device. For example, the bitstream 614 may be provided to a communication interface, network interface, wireless transmitter, port, etc. For instance, the bitstream 614 may be transmitted to another electronic device via LAN, the Internet, a cellular phone base station, etc. The bitstream 614 may additionally or alternatively be stored in memory or other component on the electronic device 602.
- [0223] Figure 6B is a block diagram illustrating one configuration of a video encoder 1782 on an electronic device 1702. The video encoder 1782 may include an enhancement layer encoder 1706, a base layer encoder 1709, a resolution upscaling block 1770 and an output interface 1780. The video encoder of Figure 6B, for example, is suitable for scalable video coding and multi-view video coding, as described herein.
- [0224] The enhancement layer encoder 1706 may include a video input 1781 that receives an input picture 1704. The output of the video input 1781 may be provided to an adder/subtractor 1783 that receives an output of a prediction selection 1750. The output of the adder/subtractor 1783 may be provided to a transform and quantize block 1752. The output of the transform and quantize block 1752 may be provided to an entropy encoding 1748 block and a scaling and inverse transform block 1772. After entropy encoding 1748 is performed, the output of the entropy encoding block 1748 may be provided to the output interface 1780. The output interface 1780 may output both the encoded base layer video bitstream 1707 and the encoded enhancement layer video bitstream 1710.
- [0225] The output of the scaling and inverse transform block 1772 may be provided to an adder 1779. The adder 1779 may also receive the output of the prediction selection 1750. The output of the adder 1779 may be provided to a deblocking block 1751. The output of the deblocking block 1751 may be provided to a reference buffer 1794. An output of the reference buffer 1794 may be provided to a motion compensation block 1754. The output of the motion compensation block 1754 may be provided to the prediction selection 1750. An output of the reference buffer 1794 may also be provided to an intra predictor 1756. The output of the intra predictor 1756 may be provided to the prediction selection 1750. The prediction selection 1750 may also receive an output of the resolution upscaling block 1770.
- [0226] The base layer encoder 1709 may include a video input 1762 that receives a downsampled input picture, or other image content suitable for combining with another image, or an alternative view input picture or the same input picture 1703 (i.e., the same as the input picture 1704 received by the enhancement layer encoder 1706). The output of the video input 1762 may be provided to an encoding prediction loop 1764. Entropy encoding 1766 may be provided on the output of the encoding prediction loop 1764. The output of the encoding prediction loop 1764 may also be provided to a

reference buffer 1768. The reference buffer 1768 may provide feedback to the encoding prediction loop 1764. The output of the reference buffer 1768 may also be provided to the resolution upscaling block 1770. Once entropy encoding 1766 has been performed, the output may be provided to the output interface 1780. The encoded base layer video bitstream 1707 and/or the encoded enhancement layer video bitstream 1710 may be provided to one or more message generation modules, as desired.

[0227] Figure 7A is a block diagram illustrating one configuration of a decoder 712 on an electronic device 702. The decoder 712 may be included in an electronic device 702. For example, the decoder 712 may be a HEVC decoder. The decoder 712 and one or more of the elements illustrated as included in the decoder 712 may be implemented in hardware, software or a combination of both. The decoder 712 may receive a bitstream 714 (e.g., one or more encoded pictures and overhead data included in the bitstream 714) for decoding. In some configurations, the received bitstream 714 may include received overhead data, such as a message (e.g., picture timing SEI message or other message), slice header, PPS, etc. In some configurations, the decoder 712 may additionally receive a separate transmission 710. The separate transmission 710 may include a message (e.g., a picture timing SEI message or other message). For example, a picture timing SEI message or other message may be received in a separate transmission 710 instead of in the bitstream 714. However, it should be noted that the separate transmission 710 may be optional and may not be utilized in some configurations.

[0228] The decoder 712 includes a CPB 720. The CPB 720 may be configured similarly to the CPB 120 described in connection with Figure 1 above. Additionally or alternatively, the decoder 712 may perform one or more of the procedures described in connection with Figure 4 and Figure 5. For example, the decoder 712 may receive a message (e.g., picture timing SEI message or other message) with sub-picture parameters and remove and decode decoding units in an access unit based on the sub-picture parameters. It should be noted that one or more access units may be included in the bitstream and may include one or more of encoded picture data and overhead data.

[0229] The Coded Picture Buffer (CPB) 720 may provide encoded picture data to an entropy decoding module 701. The encoded picture data may be entropy decoded by an entropy decoding module 701, thereby producing a motion information signal 703 and quantized, scaled and/or transformed coefficients 705.

[0230] The motion information signal 703 may be combined with a portion of a reference frame signal 798 from a decoded picture buffer 709 at a motion compensation module 780, which may produce an inter-frame prediction signal 782. The quantized, descaled and/or transformed coefficients 705 may be inverse quantized, scaled and inverse transformed by an inverse module 707, thereby producing a decoded residual signal



784. The decoded residual signal 784 may be added to a prediction signal 792 to produce a combined signal 786. The prediction signal 792 may be a signal selected from either the inter-frame prediction signal 782 produced by the motion compensation module 780 or an intra-frame prediction signal 790 produced by an intra-frame prediction module 788. In some configurations, this signal selection may be based on (e.g., controlled by) the bitstream 714.

- [0231] The intra-frame prediction signal 790 may be predicted from previously decoded information from the combined signal 786 (in the current frame, for example). The combined signal 786 may also be filtered by a de-blocking filter 794. The resulting filtered signal 796 may be written to decoded picture buffer 709. The resulting filtered signal 796 may include a decoded picture. The decoded picture buffer 709 may provide a decoded picture which may be outputted 718. In some cases 709 may be considered as frame memory.
- [0232] Figure 7B is a block diagram illustrating one configuration of a video decoder 1812 on an electronic device 1802. The video decoder 1812 may include an enhancement layer decoder 1815 and a base layer decoder 1813. The video decoder 812 may also include an interface 1889 and resolution upscaling 1870. The video decoder of Figure 7B, for example, is suitable for scalable video decoding and multi-view video decoding, as described herein.
- [0233] The interface 1889 may receive an encoded video stream 1885. The encoded video stream 1885 may consist of base layer encoded video stream and enhancement layer encoded video stream. These two streams may be sent separately or together. The interface 1889 may provide some or all of the encoded video stream 1885 to an entropy decoding block 1886 in the base layer decoder 1813. The output of the entropy decoding block 1886 may be provided to a decoding prediction loop 1887. The output of the decoding prediction loop 1887 may be provided to a reference buffer 1888. The reference buffer may provide feedback to the decoding prediction loop 1887. The reference buffer 1888 may also output the decoded base layer video stream 1884.
- [0234] The interface 1889 may also provide some or all of the encoded video stream 1885 to an entropy decoding block 1890 in the enhancement layer decoder 1815. The output of the entropy decoding block 1890 may be provided to an inverse quantization block 1891. The output of the inverse quantization block 1891 may be provided to an adder 1892. The adder 1892 may add the output of the inverse quantization block 1891 and the output of a prediction selection block 1895. The output of the adder 1892 may be provided to a deblocking block 1893. The output of the deblocking block 1893 may be provided to a reference buffer 1894. The reference buffer 1894 may output the decoded enhancement layer video stream 1882. The output of the reference buffer 1894 may also be provided to an intra predictor 1897. The enhancement layer decoder 1815 may

include motion compensation 1896. The motion compensation 1896 may be performed after the resolution upscaling 1870. The prediction selection block 1895 may receive the output of the intra predictor 1897 and the output of the motion compensation 1896. Also, the decoder may include one or more coded picture buffers, as desired, such as together with the interface 1889.

- [0235] Figure 8 illustrates various components that may be utilized in a transmitting electronic device 802. One or more of the electronic devices 102, 602, 702 described herein may be implemented in accordance with the transmitting electronic device 802 illustrated in Figure 8.
- [0236] The transmitting electronic device 802 includes a processor 817 that controls operation of the electronic device 802. The processor 817 may also be referred to as a CPU. Memory 811, which may include both read-only memory (ROM), random access memory (RAM) or any type of device that may store information, provides instructions 813a (e.g., executable instructions) and data 815a to the processor 817. A portion of the memory 811 may also include non-volatile random access memory (NVRAM). The memory 811 may be in electronic communication with the processor 817.
- [0237] Instructions 813b and data 815b may also reside in the processor 817. Instructions 813b and/or data 815b loaded into the processor 817 may also include instructions 813a and/or data 815a from memory 811 that were loaded for execution or processing by the processor 817. The instructions 813b may be executed by the processor 817 to implement the systems and methods disclosed herein. For example, the instructions 813b may be executable to perform one or more of the methods 200, 300, 400, 500 described above.
- [0238] The transmitting electronic device 802 may include one or more communication interfaces 819 for communicating with other electronic devices (e.g., receiving electronic device). The communication interfaces 819 may be based on wired communication technology, wireless communication technology, or both. Examples of a communication interface 819 include a serial port, a parallel port, a Universal Serial Bus (USB), an Ethernet adapter, an IEEE 1394 bus interface, a small computer system interface (SCSI) bus interface, an infrared (IR) communication port, a Bluetooth wireless communication adapter, a wireless transceiver in accordance with 3<sup>rd</sup> Generation Partnership Project (3GPP) specifications and so forth.
- [0239] The transmitting electronic device 802 may include one or more output devices 823 and one or more input devices 821. Examples of output devices 823 include a speaker, printer, etc. One type of output device that may be included in an electronic device 802 is a display device 825. Display devices 825 used with configurations disclosed herein may utilize any suitable image projection technology, such as a cathode ray tube (CRT), liquid crystal display (LCD), light-emitting diode (LED), gas plasma, electrolu-

minescence or the like. A display controller 827 may be provided for converting data stored in the memory 811 into text, graphics, and/or moving images (as appropriate) shown on the display 825. Examples of input devices 821 include a keyboard, mouse, microphone, remote control device, button, joystick, trackball, touchpad, touchscreen, lightpen, etc.

- [0240] The various components of the transmitting electronic device 802 are coupled together by a bus system 829, which may include a power bus, a control signal bus and a status signal bus, in addition to a data bus. However, for the sake of clarity, the various buses are illustrated in Figure 8 as the bus system 829. The transmitting electronic device 802 illustrated in Figure 8 is a functional block diagram rather than a listing of specific components.
- [0241] Figure 9 is a block diagram illustrating various components that may be utilized in a receiving electronic device 902. One or more of the electronic devices 102, 602, 702 described herein may be implemented in accordance with the receiving electronic device 902 illustrated in Figure 9.
- [0242] The receiving electronic device 902 includes a processor 917 that controls operation of the electronic device 902. The processor 917 may also be referred to as a CPU. Memory 911, which may include both read-only memory (ROM), random access memory (RAM) or any type of device that may store information, provides instructions 913a (e.g., executable instructions) and data 915a to the processor 917. A portion of the memory 911 may also include non-volatile random access memory (NVRAM). The memory 911 may be in electronic communication with the processor 917.
- [0243] Instructions 913b and data 915b may also reside in the processor 917. Instructions 913b and/or data 915b loaded into the processor 917 may also include instructions 913a and/or data 915a from memory 911 that were loaded for execution or processing by the processor 917. The instructions 913b may be executed by the processor 917 to implement the systems and methods disclosed herein. For example, the instructions 913b may be executable to perform one or more of the methods 200, 300, 400, 500 described above.
- [0244] The receiving electronic device 902 may include one or more communication interfaces 919 for communicating with other electronic devices (e.g., a transmitting electronic device). The communication interface 919 may be based on wired communication technology, wireless communication technology, or both. Examples of a communication interface 919 include a serial port, a parallel port, a Universal Serial Bus (USB), an Ethernet adapter, an IEEE 1394 bus interface, a small computer system interface (SCSI) bus interface, an infrared (IR) communication port, a Bluetooth wireless communication adapter, a wireless transceiver in accordance with 3<sup>rd</sup> Generation Partnership Project (3GPP) specifications and so forth.

- [0245] The receiving electronic device 902 may include one or more output devices 923 and one or more input devices 921. Examples of output devices 923 include a speaker, printer, etc. One type of output device that may be included in an electronic device 902 is a display device 925. Display devices 925 used with configurations disclosed herein may utilize any suitable image projection technology, such as a cathode ray tube (CRT), liquid crystal display (LCD), light-emitting diode (LED), gas plasma, electroluminescence or the like. A display controller 927 may be provided for converting data stored in the memory 911 into text, graphics, and/or moving images (as appropriate) shown on the display 925. Examples of input devices 921 include a keyboard, mouse, microphone, remote control device, button, joystick, trackball, touchpad, touchscreen, lightpen, etc.
- [0246] The various components of the receiving electronic device 902 are coupled together by a bus system 929, which may include a power bus, a control signal bus and a status signal bus, in addition to a data bus. However, for the sake of clarity, the various buses are illustrated in Figure 9 as the bus system 929. The receiving electronic device 902 illustrated in Figure 9 is a functional block diagram rather than a listing of specific components.
- [0247] Figure 10 is a block diagram illustrating one configuration of an electronic device 1002 in which systems and methods for sending a message may be implemented. The electronic device 1002 includes encoding means 1031 and transmitting means 1033. The encoding means 1031 and transmitting means 1033 may be configured to perform one or more of the functions described in connection with one or more of Figure 1, Figure 2, Figure 3, Figure 6 and Figure 8 above. For example, the encoding means 1031 and transmitting means 1033 may generate a bitstream 1014. Figure 8 above illustrates one example of a concrete apparatus structure of Figure 10. Other various structures may be implemented to realize one or more of the functions of Figure 1, Figure 2, Figure 3, Figure 6 and Figure 8. For example, a DSP may be realized by software.
- [0248] Figure 11 is a block diagram illustrating one configuration of an electronic device 1102 in which systems and methods for buffering a bitstream 1114 may be implemented. The electronic device 1102 may include receiving means 1135 and decoding means 1137. The receiving means 1135 and decoding means 1137 may be configured to perform one or more of the functions described in connection with one or more of Figure 1, Figure 4, Figure 5, Figure 7 and Figure 9 above. For example, the receiving means 1135 and decoding means 1137 may receive a bitstream 1114. Figure 9 above illustrates one example of a concrete apparatus structure of Figure 11. Other various structures may be implemented to realize one or more functions of Figure 1, Figure 4, Figure 5, Figure 7 and Figure 9. For example, a DSP may be realized by

software.

- [0249] Figure 12 is a flow diagram illustrating one configuration of a method 1200 for operation of decoded picture buffer (DPB). The method 1200 may be performed by an encoder 104 or one of its sub-parts (e.g., a decoded picture buffer module 676). The method 1200 may be performed by a decoder 112 in an electronic device 102 (e.g., electronic device B 102b). Additionally or alternatively the method 1200 may be performed by a decoder 712 or one of its sub-parts (e.g., a decoded picture buffer module 709). The decoder may parse first slice header of a picture 1202. The output and removal of pictures from DPB before decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously when first decoding unit of the access unit containing the current picture is removed from the CPB and proceeds as follows.
- [0250] - The decoding process for reference picture set (RPS) is invoked. Reference picture set is a set of reference pictures associated with a picture, consisting of all reference pictures that are prior to the associated picture in decoding order, that may be used for inter prediction of the associated picture or any picture following the associated picture in decoding order.
- [0251] - The bitstream of the video may include a syntax structure that is placed into logical data packets generally referred to as Network Abstraction Layer (NAL) units. Each NAL unit includes a NAL unit header, such as a two-byte NAL unit header (e.g., 16 bits), to identify the purpose of the associated data payload. For example, each coded slice (and/or picture) may be coded in one or more slice (and/or picture) NAL units. Other NAL units may be included for other categories of data, such as for example, supplemental enhancement information, coded slice of temporal sub-layer access (TSA) picture, coded slice of step-wise temporal sub-layer access (STSA) picture, coded slice a non-TSA, non-STSA trailing picture, coded slice of broken link access picture, coded slice of instantaneous decoded refresh picture, coded slice of clean random access picture, coded slice of decodable leading picture, coded slice of tagged for discard picture, video parameter set, sequence parameter set, picture parameter set, access unit delimiter, end of sequence, end of bitstream, filler data, and/or sequence enhancement information message. Table (4) illustrates one example of NAL unit codes and NAL unit type classes. Other NAL unit types may be included, as desired. It should also be understood that the NAL unit type values for the NAL units shown in the Table (4) may be reshuffled and reassigned. Also additional NAL unit types may be added. Also some NAL unit types may be removed.
- [0252] An intra random access point (IRAP) picture is a coded picture for which each video coding layer NAL unit has nal\_unit\_type in the range of BLA\_W\_LP to RSV\_IRAP\_VCL23, inclusive as shown in Table (4). An IRAP picture contains only

Intra coded (I) slices. An instantaneous decoding refresh (IDR) picture is an IRAP picture for which each video coding layer NAL unit has nal\_unit\_type equal to IDR\_W\_RADL or IDR\_N\_LP as shown in Table (4). An instantaneous decoding refresh (IDR) picture contains only I slices, and may be the first picture in the bitstream in decoding order, or may appear later in the bitstream. Each IDR picture is the first picture of a coded video sequence (CVS) in decoding order. A broken link access (BLA) picture is an IRAP picture for which each video coding layer NAL unit has nal\_unit\_type equal to BLA\_W\_LP, BLA\_W\_RADL, or BLA\_N\_LP as shown in Table (4). A BLA picture contains only I slices, and may be the first picture in the bitstream in decoding order, or may appear later in the bitstream. Each BLA picture begins a new coded video sequence, and has the same effect on the decoding process as an IDR picture. However, a BLA picture contains syntax elements that specify a non-empty reference picture set.

[0253]

[Table 4]

nal_unit_type	Name of nal_unit_type	Content of NAL unit and raw byte sequence payload (Rbsp) syntax structure	NAL unit type class
0 1	TRAIL_N TRAIL_R	Coded slice segment of a non-TSA, non-STSA trailing picture slice_segment_layer_rbsp()	Video Coding Layer (VCL)
2 3	TSA_N TSA_R	Coded slice segment of a temporal sub-layer access (TSA) picture slice_segment_layer_rbsp()	VCL
4 5	STSA_N STSA_R	Coded slice segment of an Step-wise Temporal sub-layer access (STSA) picture slice_segment_layer_rbsp()	VCL
6 7	RADL_N RADL_R	Coded slice segment of a random access decodable leading (RADL) picture slice_segment_layer_rbsp()	VCL
8 9	RASL_N RASL_R	Coded slice segment of a random access skipped leading (RASL) picture slice_segment_layer_rbsp()	VCL
10 12 14	RSV_VCL_N10 RSV_VCL_N12 RSV_VCL_N14	Reserved non-IRAP sub-layer non-reference VCL NAL unit types	VCL
11 13 15	RSV_VCL_R11 RSV_VCL_R13 RSV_VCL_R15	Reserved non-IRAP sub-layer reference VCL NAL unit types	VCL
16 17 18	BLA_W_LP BLA_W_RADL BLA_N_LP	Coded slice segment of a broken link access (BLA) picture slice_segment_layer_rbsp()	VCL
19 20	IDR_W_RADL IDR_N_LP	Coded slice segment of an instantaneous decoding refresh (IDR) picture slice_segment_layer_rbsp()	VCL
21	CRA_NUT	Coded slice segment of a clean random access (CRA) picture slice_segment_layer_rbsp()	VCL
22 23	RSV_IRAP_VCL22 RSV_IRAP_VCL23	Reserved IRAP VCL NAL unit types	VCL
24..31	RSV_VCL24.. RSV_VCL31	Reserved non-IRAP VCL NAL unit types	VCL
32	VPS_NUT	Video parameter set video_parameter_set_rbsp()	non-video coding layer (non-VCL)
33	SPS_NUT	Sequence parameter set seq_parameter_set_rbsp()	non-VCL
34	PPS_NUT	Picture parameter set pic_parameter_set_rbsp()	non-VCL
35	AUD_NUT	Access unit delimiter access_unit_delimiter_rbsp()	non-VCL
36	EOS_NUT	End of sequence end_of_seq_rbsp()	non-VCL
37	EOB_NUT	End of bitstream end_of_bitstream_rbsp()	non-VCL
38	FD_NUT	Filler data filler_data_rbsp()	non-VCL
39 40	PREFIX_SEI_NUT SUFFIX_SEI_NUT	Supplemental enhancement information sei_rbsp()	non-VCL
41..47	RSV_NVCL41.. RSV_NVCL47	Reserved	non-VCL
48..63	UNSPEC48.. UNSPEC63	Unspecified	non-VCL

Table (4)

[0254] Referring to Table (5), the NAL unit header syntax may include two bytes of data, namely, 16 bits. The first bit is a "forbidden\_zero\_bit" which is always set to zero at the start of a NAL unit. The next six bits is a "nal\_unit\_type" which specifies the type of raw byte sequence payloads ("Rbsp") data structure contained in the NAL unit as shown in Table (4). The next 6 bits is a "nuh\_layer\_id" which specify the identifier of the layer. In some cases these six bits may be specified as "nuh\_reserved\_zero\_6bits"

instead. The `nuh_reserved_zero_6bits` may be equal to 0 in the base specification of the standard. In a scalable video coding and/or syntax extensions `nuh_layer_id` may specify that this particular NAL unit belongs to the layer identified by the value of these 6 bits. The next syntax element is "`nuh_temporal_id_plus1`". The `nuh_temporal_id_plus1` minus 1 may specify a temporal identifier for the NAL unit. The variable temporal identifier `TemporalId` may be specified as `TemporalId = nuh_temporal_id_plus1 - 1`. The temporal identifier `TemporalId` is used to identify a temporal sub-layer. The variable `HighestTid` identifies the highest temporal sub-layer to be decoded.

[0255] [Table 5]

<code>nal_unit_header()</code> {	Descriptor
<code>forbidden_zero_bit</code>	<code>f(1)</code>
<code>nal_unit_type</code>	<code>u(6)</code>
<code>nuh_layer_id</code>	<code>u(6)</code>
<code>nuh_temporal_id_plus1</code>	<code>u(3)</code>
}	

Table (5)

[0256] Table (6) shows an exemplary sequence parameter set (SPS) syntax structure.

[0257] `pic_width_in_luma_samples` specifies the width of each decoded picture in units of luma samples. `pic_width_in_luma_samples` shall not be equal to 0.

[0258] `pic_height_in_luma_samples` specifies the height of each decoded picture in units of luma samples. `pic_height_in_luma_samples` shall not be equal to 0.

[0259] `sps_max_sub_layers_minus1` plus 1 specifies the maximum number of temporal sub-layers that may be present in each CVS referring to the SPS. The value of `sps_max_sub_layers_minus1` shall be in the range of 0 to 6, inclusive.

[0260] `sps_sub_layer_ordering_info_present_flag` flag equal to 1 specifies that `sps_max_dec_pic_buffering_minus1[i]`, `sps_max_num_reorder_pics[i]`, and `sps_max_latency_increase_plus1[i]` syntax elements are present for `sps_max_sub_layers_minus1 + 1` sub-layers.  
`sps_sub_layer_ordering_info_present_flag` equal to 0 specifies that the values of `sps_max_dec_pic_buffering_minus1[sps_max_sub_layers_minus1]`, `sps_max_num_reorder_pics[sps_max_sub_layers_minus1]`, and `sps_max_latency_increase_plus1[sps_max_sub_layers_minus1]` apply to all sub-layers.

[0261] `sps_max_dec_pic_buffering_minus1[i]` plus 1 specifies the maximum required size of the decoded picture buffer for the CVS in units of picture storage buffers when `HighestTid` is equal to `i`. The value of `sps_max_dec_pic_buffering_minus1[i]` shall be in the range of 0 to `MaxDpbSize - 1`, inclusive where `MaxDpbSize` specifies the maximum decoded picture buffer size in units of picture storage buffers. When `i` is



greater than 0,  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[i]$  shall be greater than or equal to  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[i - 1]$ . When  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[i]$  is not present for  $i$  in the range of 0 to  $\text{sps\_max\_sub\_layers\_minus1} - 1$ , inclusive, due to  $\text{sps\_sub\_layer\_ordering\_info\_present\_flag}$  being equal to 0, it is inferred to be equal to  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[\text{sps\_max\_sub\_layers\_minus1}]$ .

[0262]  $\text{sps\_max\_num\_reorder\_pics}[i]$  indicates the maximum allowed number of pictures that can precede any picture in the CVS in decoding order and follow that picture in output order when HighestTid is equal to  $i$ . The value of  $\text{sps\_max\_num\_reorder\_pics}[i]$  shall be in the range of 0 to  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[i]$ , inclusive. When  $i$  is greater than 0,  $\text{sps\_max\_num\_reorder\_pics}[i]$  shall be greater than or equal to  $\text{sps\_max\_num\_reorder\_pics}[i - 1]$ . When  $\text{sps\_max\_num\_reorder\_pics}[i]$  is not present for  $i$  in the range of 0 to  $\text{sps\_max\_sub\_layers\_minus1} - 1$ , inclusive, due to  $\text{sps\_sub\_layer\_ordering\_info\_present\_flag}$  being equal to 0, it is inferred to be equal to  $\text{sps\_max\_num\_reorder\_pics}[\text{sps\_max\_sub\_layers\_minus1}]$ .

[0263]  $\text{sps\_max\_latency\_increase\_plus1}[i]$  not equal to 0 is used to compute the value of  $\text{SpsMaxLatencyPictures}[i]$ , which specifies the maximum number of pictures that can precede any picture in the CVS in output order and follow that picture in decoding order when HighestTid is equal to  $i$ .

[0264] When  $\text{sps\_max\_latency\_increase\_plus1}[i]$  is not equal to 0, the value of  $\text{SpsMaxLatencyPictures}[i]$  is specified as follows:

[0265]  $\text{SpsMaxLatencyPictures}[i] = \text{sps\_max\_num\_reorder\_pics}[i] + \text{sps\_max\_latency\_increase\_plus1}[i] - 1$

[0266] When  $\text{sps\_max\_latency\_increase\_plus1}[i]$  is equal to 0, no corresponding limit is expressed.

[0267] The value of  $\text{sps\_max\_latency\_increase\_plus1}[i]$  shall be in the range of 0 to  $2^{32} - 2$ , inclusive. When  $\text{sps\_max\_latency\_increase\_plus1}[i]$  is not present for  $i$  in the range of 0 to  $\text{sps\_max\_sub\_layers\_minus1} - 1$ , inclusive, due to  $\text{sps\_sub\_layer\_ordering\_info\_present\_flag}$  being equal to 0, it is inferred to be equal to  $\text{sps\_max\_latency\_increase\_plus1}[\text{sps\_max\_sub\_layers\_minus1}]$ .

[0268]

[Table 6]

seq_parameter_set_rbsp() {
...
sps_max_sub_layers_minus1
...
pic_width_in_luma_samples
pic_height_in_luma_samples
...
for( i = ( sps_sub_layer_ordering_info_present_flag ? 0 :
sps_max_sub_layers_minus1 );
i <= sps_max_sub_layers_minus1; i++ ) {
sps_max_dec_pic_buffering_minus1[i]
sps_max_num_reorder_pics[i]
sps_max_latency_increase_plus1[i]
}
...
}

Table (6)

- [0269] When the current picture is an IRAP picture, the following applies:
- [0270] - If the current picture is an IDR picture, a BLA picture, the first picture in the bitstream in decoding order, or the first picture that follows an end of sequence NAL unit in decoding order, a variable NoRaslOutputFlag is set equal to 1.
- [0271] - Otherwise, if some external means is available to set a variable HandleCraAsBlaFlag to a value for the current picture, the variable HandleCraAsBlaFlag is set equal to the value provided by that external means and the variable NoRaslOutputFlag is set equal to HandleCraAsBlaFlag.
- [0272] - Otherwise, the variable HandleCraAsBlaFlag is set equal to 0 and the variable NoRaslOutputFlag is set equal to 0.
- [0273] If the current picture is an IRAP picture with NoRaslOutputFlag equal to 1 that is not picture 0, the following ordered steps are applied:
- [0274] 1. The variable NoOutputOfPriorPicsFlag is derived for the decoder under test as follows:
- [0275] - If the current picture is a CRA picture, NoOutputOfPriorPicsFlag is set equal to 1 (regardless of the value of no\_output\_of\_prior\_pics\_flag).
- [0276] - Otherwise, if the value of pic\_width\_in\_luma\_samples, pic\_height\_in\_luma\_samples, or sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] derived from the active SPS is different from the value of pic\_width\_in\_luma\_samples, pic\_height\_in\_luma\_samples, or sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ], respectively, derived from the SPS active for the preceding picture, NoOutputOfPriorPicsFlag may (but should not) be set to 1 by the decoder under test, regardless of the value of no\_output\_of\_prior\_pics\_flag.
- [0277] - Otherwise, NoOutputOfPriorPicsFlag is set equal to no\_output\_of\_prior\_pics\_flag.
- [0278] 2. The value of NoOutputOfPriorPicsFlag derived for the decoder under test is

applied for the HRD as follows:

- [0279] - If NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers in the DPB are emptied without output of the pictures they contain, and the DPB fullness is set equal to 0.
- [0280] - Otherwise (NoOutputOfPriorPicsFlag is equal to 0), all picture storage buffers containing a picture that is marked as "not needed for output" and "unused for reference" are emptied (without output), and all non-empty picture storage buffers in the DPB are emptied by repeatedly invoking the "bumping" process 1204, and the DPB fullness is set equal to 0.
- [0281] - Otherwise (the current picture is not an IRAP picture with NoRaslOutputFlag equal to 1), all picture storage buffers containing a picture which are marked as "not needed for output" and "unused for reference" are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness is decremented by one. When one or more of the following conditions are true, the "bumping" process 1204 is invoked repeatedly while further decrementing the DPB fullness by one for each additional picture storage buffer that is emptied, until none of the following conditions are true:
  - [0282] 1. The number of pictures with that particular nuh\_layer\_id value in the DPB that are marked as "needed for output" is greater than sps\_max\_num\_reorder\_pics[ HighestTid ] from the active sequence parameter set (when that particular nuh\_layer\_id value is equal to 0) or from the active layer sequence parameter set for that particular nuh\_layer\_id value.
  - [0283] 2. If sps\_max\_latency\_increase\_plus1[ HighestTid ] from the active sequence parameter set (when that particular nuh\_layer\_id value is equal to 0) or from the active layer sequence parameter set for that particular nuh\_layer\_id value is not equal to 0 and there is at least one picture with that particular nuh\_layer\_id value in the DPB that is marked as "needed for output" for which the associated variable PicLatencyCount is greater than or equal to SpsMaxLatencyPictures[ HighestTid ] for that particular nuh\_layer\_id value.
  - [0284] 3. The number of pictures with that particular nuh\_layer\_id value in the DPB is greater than or equal to sps\_max\_dec\_pic\_buffering[ HighestTid ] + 1 from the active sequence parameter set (when that particular nuh\_layer\_id value is equal to 0) or from the active layer sequence parameter set for that particular nuh\_layer\_id value.
- [0285] Picture decoding process in the block 1206 (picture decoding and marking) happens instantaneously when the last decoding unit of access unit containing the current picture is removed from the CPB.
- [0286] For each picture with nuh\_layer\_id value equal to current picture's nuh\_layer\_id value in the DPB that is marked as "needed for output", the associated variable PicLatencyCount is set equal to PicLatencyCount + 1.

- [0287] The current picture is considered as decoded after the last decoding unit of the picture is decoded. The current decoded picture is stored in an empty picture storage buffer in the DPB, and the following applies:
- [0288] - If the current decoded picture has PicOutputFlag equal to 1, it is marked as "needed for output" and its associated variable PicLatencyCount is set equal to 0.
- [0289] - Otherwise (the current decoded picture has PicOutputFlag equal to 0), it is marked as "not needed for output".
- [0290] The current decoded picture is marked as "used for short-term reference".
- [0291] When one or more of the following conditions are true, the additional "bumping" process 1208 is invoked repeatedly until none of the following conditions are true:
- [0292] - The number of pictures with nuh\_layer\_id value equal to current picture's nuh\_layer\_id value in the DPB that are marked as "needed for output" is greater than sps\_max\_num\_reorder\_pics[ HighestTid ] from the active sequence parameter set (when the current picture's nuh\_layer\_id value is equal to 0) or from the active layer sequence parameter set for the current picture's nuh\_layer\_id value.
- [0293] - sps\_max\_latency\_increase\_plus1[ HighestTid ] from the active sequence parameter set (when the current picture's nuh\_layer\_id value is equal to 0) or from the active layer sequence parameter set for the current picture's nuh\_layer\_id value is not equal to 0 and there is at least one picture with that particular nuh\_layer\_id value in the DPB that is marked as "needed for output" for which the associated variable PicLatencyCount is greater than or equal to SpsMaxLatencyPictures[ HighestTid ] for that particular nuh\_layer\_id value.
- [0294] The "bumping" process 1204 and additional bumping process 1208 are identical in terms of the steps and consists of the following ordered steps: The pictures that are first for output is selected as the ones having the smallest value of picture order count (PicOrderCntVal) of all pictures in the DPB marked as "needed for output". A picture order count is a variable that is associated with each picture, uniquely identifies the associated picture among all pictures in the CVS, and, when the associated picture is to be output from the decoded picture buffer, indicates the position of the associated picture in output order relative to the output order positions of the other pictures in the same CVS that are to be output from the decoded picture buffer.
- [0295] - These pictures are cropped, using the conformance cropping window specified in the active sequence parameter set for the picture with nuh\_layer\_id equal to 0 or in the active layer sequence parameter set for a nuh\_layer\_id value equal to that of the picture, the cropped pictures are output in ascending order of nuh\_layer\_id, and the pictures are marked as "not needed for output".
- [0296] - Each picture storage buffer that contains a picture marked as "unused for reference" and that included one of the pictures that was cropped and output is emptied.

- [0297] Referring to Figure 13A, as previously described the NAL unit header syntax may include two bytes of data, namely, 16 bits. The first bit is a "forbidden\_zero\_bit" which is always set to zero at the start of a NAL unit. The next six bits is a "nal\_unit\_type" which specifies the type of raw byte sequence payloads ("RBSP") data structure contained in the NAL unit. The next 6 bits is a "nuh\_reserved\_zero\_6bits". The nuh\_reserved\_zero\_6bits may be equal to 0 in the base specification of the standard. Other values of nuh\_reserved\_zero\_6bits may be specified as desired. Decoders may ignore (i.e., remove from the bitstream and discard) all NAL units with values of nuh\_reserved\_zero\_6bits not equal to 0 when handling a stream based on the base specification of the standard. In a scalable or other extension nuh\_reserved\_zero\_6bits may specify other values, to signal scalable video coding and/or syntax extensions. In some cases syntax element nuh\_reserved\_zero\_6bits may be called reserved\_zero\_6bits. In some cases the syntax element nuh\_reserved\_zero\_6bits may be called as layer\_id\_plus1 or layer\_id, as illustrated in Figure 13B and Figure 13C. In this case the element layer\_id will be layer\_id\_plus1 minus 1. In this case it may be used to signal information related to layer of scalable coded video. The next syntax element is "nuh\_temporal\_id\_plus1". nuh\_temporal\_id\_plus1 minus 1 may specify a temporal identifier for the NAL unit. The variable temporal identifier TemporalId may be specified as  $\text{TemporalId} = \text{nuh\_temporal\_id\_plus1} - 1$ .
- [0298] Referring to Figure 14, a general NAL unit syntax structure is illustrated. The NAL unit header two byte syntax of Figure 13 is included in the reference to nal\_unit\_header() of Figure 14. The remainder of the NAL unit syntax primarily relates to the RBSP.
- [0299] One existing technique for using the "nuh\_reserved\_zero\_6bits" is to signal scalable video coding information by partitioning the 6 bits of the nuh\_reserved\_zero\_6bits into distinct bit fields, namely, one or more of a dependency ID, a quality ID, a view ID, and a depth flag, each of which refers to the identification of a different layer of the scalable coded video. Accordingly, the 6 bits indicate what layer of the scalable encoding technique this particular NAL unit belongs to. Then in a data payload, such as a video parameter set ("VPS") extension syntax ("scalability\_type") as illustrated in Figure 15, the information about the layer is defined. The VPS extension syntax of Figure 15 includes 4 bits for scalability type (syntax element scalability\_type) which specifies the scalability types in use in the coded video sequence and the dimensions signaled through layer\_id\_plus1 (or layer\_id) in the NAL unit header. When the scalability type is equal to 0, the coded video sequence conforms to the base specification, thus layer\_id\_plus1 of all NAL units is equal to 0 and there are no NAL units belonging to an enhancement layer or view. Higher values of the scalability type are interpreted as illustrated in Figure 16.

- [0300] The `layer_id_dim_len[ i ]` specifies the length, in bits, of the *i*-th scalability dimension ID. The sum of the values `layer_id_dim_len[ i ]` for all *i* values in the range of 0 to 7 is less than or equal to 6. The `vps_extension_byte_alignment_reserved_zero_bit` is zero. The `vps_layer_id[ i ]` specifies the value of `layer_id` of the *i*-th layer to which the following layer dependency information applies. The `num_direct_ref_layers[ i ]` specifies the number of layers the *i*-th layer directly depends on. The `ref_layer_id[ i ][ j ]` identifies the *j*-th layer the *i*-th layer directly depends on.
- [0301] In this manner, the existing technique signals the scalability identifiers in the NAL unit and in the video parameter set to allocate the bits among the scalability types listed in Figure 16. Then for each scalability type, Figure 16 defines how many dimensions are supported. For example, scalability type 1 has 2 dimensions (i.e., spatial and quality). For each of the dimensions, the `layer_id_dim_len[i]` defines the number of bits allocated to each of these two dimensions, where the total sum of all the values of `layer_id_dim_len[i]` is less than or equal to 6, which is the number of bits in the `nuh_reserved_zero_6bits` of the NAL unit header. Thus, in combination the technique identifies which types of scalability is in use and how the 6 bits of the NAL unit header are allocated among the scalability.
- [0302] While such a fixed combination of different scalability dimensions, as illustrated in Figure 16, is suitable for many applications there are desirable combinations which are not included. Referring to Figure 17, a modified video parameter set extension syntax specifies a scalability type for each bit in the `nuh_reserved_zero_6bits` syntax element. The `vps_extension_byte_alignment_reserved_zero_bit` is set to 0. The `max_num_layers_minus1_bits` indicates the total number of bits used for the syntax element in the first two bytes of the NAL unit header in Figure 13 referred to as `layer_id_plus1` or `nuh_reserved_zero_6bits`. The `scalability_map[i]` specifies the scalability type for each bit in the `layer_id_plus1` syntax element. In some case the `layer_id_plus1` syntax element may be instead called `nuh_reserved_zero_6bits` or `rserved_zero_6bits` syntax element. The scalability map for all the bits of the syntax element `layer_id_plus1` together specifies the scalability in use in the coded video sequence. The actual value of the identifier for each of the scalability types is signaled through those corresponding bits in the `layer_id_plus1` (`nuh_reserved_zero_6bits`) field in the NAL unit header. When `scalability_map[i]` is equal to 0 for all values of *i*, the coded video sequence conforms to the base specification, thus `layer_id_plus1` value of NAL units is equal to 0 and there are no NAL units belonging to an enhancement layer or view. The `vps_layer_id[ i ]` specifies the value of `layer_id` of the *i*-th layer to which the following layer dependency information applies. The `num_direct_ref_layers[ i ]` specifies the number of layers the *i*-th layer directly depends on. The `ref_layer_id[ i ][ j ]`

- ] identifies the j-th layer the i-th layer directly depends on.
- [0303] Higher values of scalability\_map[i] are interpreted as shown in Figure 18. The scalability map [i] includes the scalability dimensions of (0) none; (1) spatial; (2) quality; (3) depth; (4) multiview; (5) unspecified; (6) reserved; and (7) reserved.
- [0304] Therefore each bit in the NAL unit header is interpreted based on the 3 bits in the video parameter set of what is the scalability dimension (e.g., none, spatial, quality, depth, multiview, unspecified, reserved). For example, to signal that all the bits in layer\_id\_plus1 correspond to spatial scalability, the scalability\_map values in the VPS may be coded as 001 001 001 001 001 001 for the 6 bits of the NAL unit header. Also for example, to signal that 3 bits in layer\_id\_plus1 correspond to spatial scalability and 3 bits correspond to quality scalability, the scalability\_map values in the VPS may be coded as 001 001 001 010 010 010 for the 6 bits of the NAL Unit header.
- [0305] Referring to Figure 19, another embodiment includes the video parameter set signaling the number of scalability dimensions in the 6 bits of the NAL unit header using the num\_scalability\_dimensions\_minus1. The num\_scalability\_dimensions\_minus1 plus 1 indicates the number of scalability dimensions signaled through the layer\_id\_plus1; nuh\_reserved\_zero\_6bits; and/or reserved\_zero\_6bits syntax elements. The scalability\_map[i] has the same semantics as described above in relation to Figure 17. The num\_bits\_for\_scalability\_map[i] specifies the length in bits for the i'th scalability dimension. The sum of all of the num\_bits\_for\_scalability\_map[i] for i=0, ..num\_scalability\_dimensions\_minus1 is equal to six (or otherwise equal to the number of bits used for layer\_id\_plus1; vps\_reserved\_zero\_6bits; max\_num\_layers\_minus1; reserved\_zero\_6bits; nuh\_reserved\_zero\_6bits syntax elements).
- [0306] With respect to Figure 17 and Figure 19 other variations may be used, if desired. In one embodiment for example, the scalability\_map[i] may be signaled with u(4) (or u(n) with n>3 or n<3). In this case the higher values of scalability\_map[i] may be specified as reserved for bitstreams conforming to a particular profile of the video technique. For example, scalability map values 6..15 may be specified as 'reserved' when signaling scalability\_map[i] with u(4). In another embodiment for example, scalability\_map[i] maybe signaled with ue(v) or some other coding scheme. In another embodiment for example, a restriction may be specified such that the scalability\_map[i] values are arranged in monotonic non decreasing (or non-increasing) order. This results in various scalability dimension fields in the layer\_id\_plus1 field in NAL unit header being contiguous.
- [0307] Another existing technique for signaling the scalable video coding using the "layer\_id\_plus1" or "nuh\_reserved\_zero\_6bits" syntax element is to map the layer\_id\_plus1 in the NAL unit header to a layer identification by signaling a general

lookup table in the video parameter set. Referring to Figure 20, the existing technique includes a video parameter set that specifies the number of dimension types and dimension identifications for the i-th layer of the lookup table. In particular, the `vps_extension_byte_alignment_reserved_zero_bit` is zero. The `num_dimensions_minus1[i]` plus 1 specifies the number of dimension types (`dimension_type[i][j]`) and dimension identifiers (`dimension_id[i][j]`) for the i-th layer. The `dimension_type[i][j]` specifies the j-th scalability dimension type of the i-th layer, which has `layer_id` or `layer_id_plus1` equal to i, as specified in Figure 31. As illustrated in Figure 21, the dimensions that are identified include of (0) view order idx; (1) depth flag; (2) dependency ID; (3) quality ID; (4)-(15) reserved. The `dimension_id[i][j]` specifies the identifier of the j-th scalability dimension type of the i-th layer, which when not present is inferred to be 0. The `num_direct_ref_layers[i]` specifies the number of layers the i-th layer directly depends on. The `ref_layer_id[i][j]` identifies the j-th layer the i-th layer directly depends on. Unfortunately, the proposed embodiment illustrated in Figure 20 results in an unwieldy large lookup table.

[0308] Referring to Figure 22, a modified video parameter set extension includes a scalability mask that is used in combination with a scalability dimension. The `scalability_mask` signals a pattern of 0 and 1 bits with each bit corresponding to one scalability dimension as indicated by the scalability map syntax of Figure 23. A value of 1 for a particular scalability dimension indicates that this scalability dimension is present in this layer (I'th layer). A value of 0 for a particular scalability dimension indicates that this scalability dimension is not present in this layer (I'th layer). For example, a set of bits of 00100000 refers to quality scalability. The actual identifier value of the particular scalability dimension that is present is indicated by the `scalability_id[j]` value signaled. The values of `num_scalability_types[i]` is equal to the sum of number of bits in the `scalability_mask` having value of 1. Thus

[0309] [Math.9]

$$\text{num\_scalability\_types}[i] = \sum_{k=0}^7 \text{scalability\_mask}[i](k)$$

[0310] The `scalability_id[j]` indicates the j-th scalability dimension's identifier value for the type of scalability values that are signaled by the `scalability_mask` value.

[0311] Referring to Figure 24, a modification of Figure 22, includes the scalability mask being signaled outside the loop. This results in one common mask for each layer identification. Referring to Figure 25, in this modification a corresponding exemplary video parameter set may include the scalable identification with the scalability mask not being included. In this case the syntax element `scalable_id[j]` has same interpretation as the syntax element `scalability_id[j]` in Figure 22.

[0312] Referring to Figure 26 a modification of Figure 22 includes the scalability mask



(scalability\_mask) being signaled outside the loop. This results in one common mask for each layer identification. The scalability\_mask signals a pattern of 0 and 1 bits with each bit corresponding to one scalability dimension as indicated by the scalability map syntax of Figure 27. A value of 1 for a particular scalability dimension indicates that this scalability dimension is present in this layer (i'th layer). A value of 0 for a particular scalability dimension indicates that this scalability dimension is not present in this layer (i'th layer). For example, a set of bits of 00100000 refers to quality scalability. The actual identifier value of the particular scalability dimension that is present is indicated by the scalability\_id[j] value signaled. The values of num\_scalability\_types[i] is equal to the sum of number of bits in the scalability\_mask having value of 1. Thus

[0313] [Math.10]

$$\text{NumScalabilityTypes}[i] = \sum_{k=0}^{15} \text{scalability\_mask}(k)$$

[0314] In this case the scalability\_id[j] variable may instead be called dimension\_id[i][j] variable. dimension\_id[i][j] specifies the scalability identifier of the j-th scalability dimension of the i-th layer. Then a variable ScalabilityId[i][j] is derived as follows.

[0315] [Table 7]

for( i = 1; i <= vps_max_layers_minus1; i++ ) {
for(k=0, j=0; k<=15; k++) {
if(scalability_mask(k)==1)
ScalabilityId [i][k]=dimension_id[i][j++]
else
ScalabilityId [i][k]=0;
}
}

[0316] Where the ScalabilityId [i][k] signals dimension ID for the corresponding scalability type as follows.

[0317] [Table 8]

k	ScalabilityId [i][k]
0	DependencyId[i][k]
1	QualityId[i][k]
2	depthFlag[i][k]
3	ViewId[i][k]
4-15	Reserved

[0318] Where DependencyId[i][1] is the dependency ID for the spatial scalability dimension for the i-th layer, QualityId[i][2] is the quality ID for the quality scalability dimension for the i-th layer, depthFlag[i][3] is the depth flag / depth ID for the depth scalability

dimension for the i-th layer, and ViewId[i][4] is the view ID for the multiview scalability dimension for the i-th layer.

[0319] Also in Figure 26 `avc_base_codec_flag` equal to 1 specifies that the base layer conforms to Rec. ITU-T H.264 | ISO/IEC 14496-10, and `avc_base_codec_flag` equal to 1 specifies to HEVC. `vps_nuh_layer_id_presnet_flag` indicates if `layer_id_in_nuh[i]` variable which signals the value of `layer_id` in NAL unit header is signaled.

[0320] In another embodiment one or more of the syntax elements `scalability_mask[i]`, `scalability_mask`, `scalability_id[j]` may be signaled using different number of bits than `u(8)`. For example they could be signaled with `u(16)` (or `u(n)` with  $n > 8$  or  $n < 8$ ). In another embodiment one or more of these syntax element could be signaled with `ue(v)`. In another embodiment the `scalability_mask` may be signaled in the NAL unit header in `layer_id_plus1`; `vps_reserved_zero_6bits`; `max_num_layers_minus1`; `reserved_zero_6bits`; and/or `nuh_reserved_zero_6bits` syntax elements. In some embodiments the system may do this only for VPS NAL units, or only for non-VPS NAL units, or for all NAL units. In yet another embodiment `scalability_mask` may be signaled per picture anywhere in the bitstream. For example it may be signaled in slice header, picture parameter set, video parameter set, or any other parameter set or any other normative part of the bitstream.

[0321] It should be noted that FIGS. 13, 15, 18, 20, 21, 22, 23 and corresponding description refer to 6 bits since the syntax element `nuh_reserved_zero_6bits` or `layer_id_plus1` in NAL unit header of Figure 13 has 6 bits. However all the above description can be suitably modified if that syntax element used a different number of bits than 6 bits. For example if that syntax element (`nuh_reserved_zero_6bits` or `layer_id_plus1`) instead used 9 bits then in Figure 17 the value of `max_num_layer_minus1` bits will be 9 and the `scalability_map[i]` will be signaled for each of the 9 bits instead of 6 bits.

[0322] Referring to FIG 24 a modification of FIG 22 provides syntax for signaling layer dependency information. New syntax element `layer_dependency_information_pattern` is defined.

[0323] `layer_dependency_information_pattern` signals a pattern of 0 and 1 bits with the length equal to `vps_max_layers_minus1`. A value of 0 for i'th bit indicates that the layer with `layer_id (i+1)` is an independent layer. A value of 1 for i'th bit indicates that the layer with `layer_id (i+1)` is a dependent layer which depends on one or more of other layers.

[0324] The values of `NumDepLayers` is equal to the sum of number of bits in the `layer_dependency_information_pattern` having value of 1. Thus

[0325] [Math.11]

$$\text{NumDepLayers} = \sum_{k=0}^{\text{vps\_max\_layer\_minus1}-1} \text{layer\_dependency\_information\_pattern}(k)$$

- [0326] Referring to Figure 29 a modification of Figure 26 provides syntax for signaling layer dependency information. New syntax element `layer_dependency_flag[i]` is defined. `layer_dependency_flag[i]` signals if a layer depends on other layers. A value of 0 for the flag indicates that the layer with `layer_id i` is an independent layer. A value of 1 for `i`'th bit indicates that the layer with `layer_id i` is a dependent layer.
- [0327] Referring to Figure 30 a modification of Figure 26 provides syntax for signaling layer dependency information. New syntax element `layer_dependency_map[i]` is defined. `layer_dependency_map[i]` signals a pattern of 0 and 1 bits with the length equal to `vps_max_layers_minus1`. A value of 0 for `k`'th bit of `layer_dependency_map[i]` indicates that the layer `i` does not depend on layer with `layer_id (k+1)`. A value of 1 for `k`'th bit of `layer_dependency_map[i]` indicates that the layer `i` depends on layer with `layer_id (k+1)`.
- [0328] Referring to Figure 31 a modification of Figure 26 provides syntax for signaling layer dependency information. New syntax element `layer_dependency_information_pattern` is defined.
- [0329] `layer_dependency_information_pattern` signals a pattern of 0 and 1 bits with the length equal to `vps_max_layers_minus1`. A value of 0 for `i`'th bit indicates that the layer with `layer_id (i+1)` is an independent layer. A value of 1 for `i`'th bit indicates that the layer with `layer_id (i+1)` is a dependent layer which depends on one or more of other layers. The values of `NumDepLayers` is equal to the sum of number of bits in the `layer_dependency_information_pattern` having value of 1. Thus
- [0330] [Math.12]

$$\text{NumDepLayers} = \sum_{k=0}^{\text{vps\_max\_layer\_minus1}-1} \text{layer\_dependency\_information\_pattern}(k)$$

- [0331] `layer_dependency_map[i]` signals a pattern of 0 and 1 bits with the length equal to `vps_max_layers_minus1`. A value of 0 for `k`'th bit of `layer_dependency_map[i]` indicates that the layer `i` does not depend on layer with `layer_id (k+1)`. A value of 1 for `k`'th bit of `layer_dependency_map[i]` indicates that the layer `i` depends on layer with `layer_id (k+1)`.
- [0332] Referring to Figure 32 a modification of Figure 26 provides syntax for signaling layer dependency information. FIG 28 is a variant syntax based on syntax in FIG 27. New syntax element `layer_dependency_information_pattern` is defined.
- [0333] `layer_dependency_information_pattern` signals a pattern of 0 and 1 bits with the length equal to `vps_max_layers_minus1`. A value of 0 for `i`'th bit indicates that the layer with `layer_id (i+1)` is an independent layer. A value of 1 for `i`'th bit indicates that the layer with `layer_id (i+1)` is a dependent layer which depends on one or more of other layers.

[0334] The values of NumDepLayers is equal to the sum of number of bits in the layer\_dependency\_information\_pattern having value of 1. Thus

[0335] [Math.13]

$$\text{NumDepLayers} = \sum_{k=0}^{\text{vps\_max\_layer\_minus1}-1} \text{layer\_dependency\_information\_pattern}(k)$$

[0336] Syntax elements num\_direct\_ref\_layers[i] and ref\_layer\_id[i][j] are signaled only when layer\_dependency\_information\_pattern(i) has a value of 1. Where layer\_dependency\_information\_pattern(i) is the i'th bit of the syntax element layer\_dependency\_pattern.

[0337] Referring to Figure 33 a modification of Figure 26 provides syntax for signaling layer dependency information. FIG 29 is a variant syntax based on syntax in Figure 31. New syntax element layer\_dependency\_information\_pattern is defined.

[0338] layer\_dependency\_information\_pattern signals a pattern of 0 and 1 bits with the length equal to vps\_max\_layers\_minus1. A value of 0 for i'th bit indicates that the layer with layer\_id (i+1) is an independent layer. A value of 1 for i'th bit indicates that the layer with layer\_id (i+1) is a dependent layer which depends on one or more of other layers.

[0339] The values of NumDepLayers is equal to the sum of number of bits in the layer\_dependency\_information\_pattern having value of 1. Thus

[0340] [Math.14]

$$\text{NumDepLayers} = \sum_{k=0}^{\text{vps\_max\_layer\_minus1}-1} \text{layer\_dependency\_information\_pattern}(k)$$

[0341] layer\_dependency\_map[i] signals a pattern of 0 and 1 bits with the length equal to vps\_max\_layers\_minus1. A value of 0 for k'th bit of layer\_dependency\_map[i] indicates that the layer i does not depend on layer with layer\_id (k+1). A value of 1 for k'th bit of layer\_dependency\_map[i] indicates that the layer i depends on layer with layer\_id (k+1). Syntax elements layer\_dependency\_map[i] is signaled only when layer\_dependency\_information\_pattern(i) has a value of 1. Where layer\_dependency\_information\_pattern(i) is the i'th bit of the syntax element layer\_dependency\_pattern.

[0342] In another embodiment layer\_dependency\_information\_pattern syntax element may be signaled as a set of 1 bit flag values. In this case a total of vps\_max\_layers\_minus1 1 bit values will be signaled as:

[0343]

[Table 9]

for( i = 1; i <= vps_max_layers_minus1 ; i++ )
{
layer_dependency_information_pattern_flags[i];
}

[0344] In another embodiment layer\_dependency\_map[i] syntax element may be signaled as a set of 1 bit flag values. In this case a total of vps\_max\_layers\_minus1 1 bit values will be signaled as:

[0345] [Table 10]

for( j = 1; j <= vps_max_layers_minus1 ; j++ )
{
layer_dependency_map_values[i][j];
}

[0346] In another embodiment one or more of the syntax elements layer\_dependency\_information\_pattern, layer\_dependency\_map may be signaled using a known fixed number of bits instead of u(v). For example they could be signaled using u(64).

[0347] In another embodiment one or more of or more of the syntax elements layer\_dependency\_information\_pattern, layer\_dependency\_map may be signaled with ue(v) or some other coding scheme.

[0348] In another embodiment the names of various syntax elements and their semantics may be altered by adding a plus1 or plus2 or by subtracting a minus1 or a minus2 compared to the described syntax and semantics.

[0349] In yet another embodiment various syntax elements such as layer\_dependency\_information\_pattern, layer\_dependency\_map, layer\_dependency\_flag[i] etc. may be signaled per picture anywhere in the bitstream. For example it may be signaled in slice header, pps/ sps/ vps/ aps or any other parameter set or other normative part of the bitstream.

[0350] As previously described, scalable video coding is a technique of encoding a video bitstream that also contains one or more subset bitstreams. A subset video bitstream may be derived by dropping packets from the larger video to reduce the bandwidth required for the subset bitstream. The subset bitstream may represent a lower spatial resolution (smaller screen), lower temporal resolution (lower frame rate), or lower quality video signal. For example, a video bitstream may include 5 subset bitstreams, where each of the subset bitstreams adds additional content to a base bitstream.

Hannuksela, et al., "Test Model for Scalable Extensions of High Efficiency Video Coding (HEVC)" JCTVC-L0453, Shanghai, October 2012, Chen, et al., "SHVC Draft Text 1," JCTVC-L1008, Geneva, March, 2013, and Wang, et al., "AHG9: On VPS and SPS in HEVC 3DV and scalable extensions," JCTVC-M0268, Incheon, April 2013 are referenced herein.

[0351] As previously described, multi-view video coding is a technique of encoding a video bitstream that also contains one or more other bitstreams representative of alternative views. For example, the multiple views may be a pair of views for stereoscopic video. For example, the multiple views may represent multiple views of the same scene from different viewpoints. The multiple views generally contain a large amount of inter-view statistical dependencies, since the images are of the same scene from different viewpoints. Therefore, combined temporal and inter-view prediction may achieve efficient multi-view encoding. For example, a frame may be efficiently predicted not only from temporally related frames, but also from the frames of neighboring viewpoints. Hannuksela, et al., "Common specification text for scalable and multi-view extensions," JCTVC-L0452, Geneva, January 2013, and Tech, et. al. "MV-HEVC Draft Text 3 (ISO/IEC 23008-2:201x/PDAM2)," JCT3V-C1004\_d3, Geneva, January 2013, are referenced herein.

[0352] Chen, et al., "SHVC Draft Text 1," JCTVC-L1008, Geneva, January 2013; Hannuksela, et al. "Test Model for Scalable Extensions of High Efficiency Video Coding (HEVC)," JCTVC-L0453-spec-text, Shanghai, October 2012; and Hannuksela, "Draft Text for Multiview Extension of High Efficiency Video Coding (HEVC)," JCTVC-L0452-spec-text-r1, Shanghai, October 2012 are referenced herein. JCTVC-L0452 and JCTVC-L0453 each have the concept of output operation points. This was later changed to the concept of output layer sets in JCTVC-L1008. For the syntax elements `num_output_operation_points`, `output_op_point_index[]`, and `output_layer_flag[][]` are defined for the output operation points for JCTVC-L0453 and LCTVC-L0452, and modified to `num_output_layer_sets`, `output_layer_set_idx[]`, and `output_layer_flag[][]` in JCTVC-L1008. In each of such documents, a layer dependency change SEI message is defined. The SEI message allows signaling layer dependency information changes starting with the current access unit containing the SEI message.

[0353] Chen, et al., "SHVC Draft Text 1," JCTVC-L1008, Geneva, January 2013; Hannuksela, et al. "Test Model for Scalable Extensions of High Efficiency Video Coding (HEVC)," JCTVC-L0453-spec-text, Shanghai, October 2012; and Hannuksela, "Draft Text for Multiview Extension of High Efficiency Video Coding (HEVC),"

JCTVC-L0452-spec-text-r1, Shanghai, October 2012; each have an output order decoded picture buffer (DPB) which operates based on using `sps_max_num_reorder_pics[HighestTid]`, `sps_max_latency_increase_plus1[HighestTid]` and `sps_max_dec_pic_buffering[HighestTid]` syntax elements for the output and removal of pictures 0 from the DPB. This information is signaled in the video parameter set for the base layer, which provides buffering information for the video content including the enhancement layers, if any.

- [0354] Referring to Figure 34, another example of a video parameter set is illustrated. In particular, the video parameter set of Figure 34 includes reference to an associated video parameter set extension (`vps_extension`). The `vps_extension` may be signaled with the `vps_extension_flag`, the `vps_extension2_flag`, and/or the `vps_extension_data_flag`.
- [0355] Referring to Figure 35, another example of a video parameter set extension is illustrated. In particular, the video parameter set extension of Figure 35 includes reference to `num_output_layer_sets`, which defines those layers of a set of layers of a bitstream that may be output to the viewer by the decoder. `num_output_layer_sets` specifies the number of layer sets for which output layers are specified with `output_layer_set_index[ i ]` and `output_layer_flag[ lIdx ][ j ]`. When not present, the value of `num_output_layer_sets` is inferred to be equal to 0. A layer set describing output layers is an output layer set.
- [0356] The `output_layer_set_idx[ i ]` specifies the index `lIdx` of the layer set for which `output_layer_flag[ lIdx ][ j ]` is present.
- [0357] The `output_layer_flag[ lIdx ][ j ]` equal to 1 specifies that the layer with `nuh_layer_id` equal to `j` is a target output layer of the `lIdx`-th layer set. A value of `output_layer_flag[ lIdx ][ j ]` equal to 0 specifies that the layer with `nuh_layer_id` equal to `j` is not a target output layer of the `lIdx`-th layer set.
- [0358] With the `num_output_layer_sets` being defined in the `vps_extension` in the manner described, the `num_output_layer_sets` may be updated by sending a new video parameter set together with a corresponding video parameter set extension. Unfortunately, sending the new video parameter set together with the corresponding video parameter set extension results in a significant reduction in the available bandwidth. Additionally a new video parameter set could only be activated at certain picture types, for example at Intra random access point picture types.
- [0359] As an example, a bitstream may include a base layer 0 and four enhancement layers, namely enhancement layer 1, enhancement layer 2, enhancement layer 3, and enhancement layer 4. A first layer set may be base layer 0 and enhancement layer 1. A second layer set may be base layer 0, enhancement layer 1, and enhancement layer 2.

A third layer set may be base layer 0, enhancement layer 1, and enhancement layer 3. A fourth layer set may be base layer 0 and enhancement layer 4. The output layer sets define the particular layers of a layer set that may be provided as an output. For example, output layer set 1 may be layer 0 is an output and layer 1 is not an output. For example, output layer set 2 may be layer 0 is an output, layer 1 is an output, and layer 2 is not output. For example, output layer set 3 may be layer 0 is not output, layer 1 is not output, and layer 3 is output. For example, output layer set 4 may be layer 0 is output and layer 4 is not output. A limited number of output layers is useful to accommodate limited decoding capabilities. In some cases, it is desirable to disable some dependencies between different layers to accommodate decoding other layers and/or decoder capabilities. When some layer dependencies change it may likewise be desirable to enable, disable, and/or add other layers as output layers. As an example layer dependency change could require fewer layers to be decoded to decode a particular target layer (e.g. a view in MV-HEVC). Thus a decoder could decode and output a layer (e.g. and additional view) after a layer dependency change that it previously could not due to its hardware limitation. Currently no mechanism exists to allow signaling a change in the output layer sets without sending a new video parameter set.

[0360] Referring to Figure 36, a SEI message syntax may be used to facilitate changes of the output layer sets, namely, `output_layer_sets_change(payloadSize)`. This SEI message indicates that the output layer sets information changes starting with the current access unit containing the SEI message and is interpreted with respect to the active video parameter set. More than one video parameter set may be included with the bitstream, and the active video parameter set is changed, as desired. When present, the output layer sets change SEI message applies to the target access unit set that consists of the current access unit and all the subsequent access units, in decoding order, until the next output layer sets change SEI message or the end of the coded video sequence, whichever is earlier in decoding order. Output layer sets change SEI messages preferably do not have a cumulative effect.

[0361] The 'active\_vps\_id' identifies an active video parameter set that contains the output layer sets information to which the change or addition applied. The value of `active_vps_id` is equal to the value of `video_parameter_set_id` of the active video parameter set for the VCL NAL units of the access unit containing the SEI message.

[0362] The 'num\_changed\_output\_layer\_sets' specifies the number of changed output layer sets for which output layers are specified in vps extension section of the active video parameter set identified by `active_vps_id`. The value of `num_changed_output_layer_sets` should be in the range of 0 to `num_output_layer_sets`, inclusive. When not present, the value of `num_changed_output_layer_sets` is inferred



to be equal to 0.

- [0363] The 'changed\_output\_layer\_sets\_idx\_entry'[ i ] identifies the index entry in the list of output layer set entries in the vps extension identified by active\_vps\_id for which output\_layer\_flag[ clsIdx ][ j ] change applies to. The value of changed\_output\_layer\_set\_idx\_entry[ i ] should be in the range of 0 to num\_output\_layer\_sets, inclusive. The changed\_output\_layer\_sets\_idx\_entry[ i ] is not equal to changed\_output\_layer\_sets\_idx\_entry[ j ] for any j not equal to i.
- [0364] The 'output\_layer\_flag'[ clsIdx ][ j ] equal to 1 specifies that the layer with nuh\_layer\_id equal to j is a target output layer of the clsIdx-th layer set. A value of output\_layer\_flag[ clsIdx ][ j ] equal to 0 specifies that the layer with nuh\_layer\_id equal to j is not a target output layer of the clsIdx-th layer set.
- [0365] When output\_layer\_flag[ clsIdx ][ j ] is not present for clsIdx in the range of 0 to vps\_num\_layer\_sets\_minus1, inclusive and for j in the range of 0 to 63, inclusive, output\_layer\_flag[ clsIdx ][ j ] is inferred to be equal to ( j == LayerSetLayerIdList[ clsIdx ][ NumLayersInIdList[ clsIdx ] - 1 ] ).
- [0366] The 'num\_addl\_output\_layer\_sets' specifies the number of additional layer sets for which output layers are specified with addl\_output\_layer\_set\_idx[ i ] and output\_layer\_flag[ addlIdx ][ j ]. When not present, the value of num\_addl\_output\_layer\_sets is inferred to be equal to 0. The 'addl\_output\_layer\_sets\_idx'[ i ] identifies the index addlIdx of the layer set for which output\_layer\_flag[addlIdx][j] is present.
- [0367] The 'output\_layer\_flag'[ addlIdx ][ j ] equal to 1 specifies that the layer with nuh\_layer\_id equal to j is a target output layer of the addlIdx-th layer set. A value of output\_layer\_flag[ addlIdx ][ j ] equal to 0 specifies that the layer with nuh\_layer\_id equal to j is not a target output layer of the addlIdx-th layer set.
- [0368] When output\_layer\_flag[ addlIdx ][ j ] is not present for addlIdx in the range of 0 to vps\_num\_layer\_sets\_minus1, inclusive and for j in the range of 0 to 63, inclusive, output\_layer\_flag[ addlIdx ][ j ] is inferred to be equal to ( j == LayerSetLayerIdList[ addlIdx ][ NumLayersInIdList[ addlIdx ] - 1 ] ).
- [0369] When output\_layer\_flag[ addlIdx ][ j ] is not present for addlIdx in the range of 0 to vps\_num\_layer\_sets\_minus1, inclusive and for j in the range of 0 to 63, inclusive, and when output\_layer\_flag[ clsIdx ][ j ] is not present for clsIdx in the range of 0 to vps\_num\_layer\_sets\_minus1, inclusive and for j in the range of 0 to 63, inclusive, output\_layer\_flag[ addlIdx ][ j ] is inferred to be equal to ( j == LayerSetLayerIdList[ addlIdx ][ NumLayersInIdList[ addlIdx ] - 1 ] ).
- [0370] When output\_layer\_flag[ addlIdx ][ j ] is not present for addlIdx in the range of 0 to vps\_num\_layer\_sets\_minus1, inclusive and for j in the range of 0 to 63, inclusive, and when output\_layer\_flag[ clsIdx ][ j ] is not present for clsIdx in the range of 0 to

vps\_num\_layer\_sets\_minus1, inclusive and for j in the range of 0 to 63, inclusive, output\_layer\_flag[ clsIdx ][ j ] is inferred to be equal to ( j == LayerSetLayerIdList[ clsIdx ][ NumLayersInIdList[ clsIdx ] - 1 ] ).

[0371] A target list of output layers may be derived, if desired. The target output layer identifier list, targetOpLayerIdList, specifies the list of nuh\_layer\_id values, in increasing order of nuh\_layer\_id values which are output layers for the selected output layer set may be derived as follows:

[0372] For a selected output layer set oLsIdx, a target output layer identifier list targetOpLayerIdList is derived as:

[0373] [Table 11]

for(k=0, numOutputLayers=0; k<=vps_max_layer_id; k++)
if(output_layer_flag[oLsIdx][k])
targetOpLayerIdList[numOutputLayers++] = layer_id_in_nuh[k]

[0374] A target output layer set may be identified by the associated target output layer identifier list targetOpLayerIdList.

[0375] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets if there are no output layer sets change SEI messages.

[0376] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets + num\_addl\_output\_layer\_sets if there are output layer sets change SEI messages present.

[0377] A target list of decoded layers may be derived, if desired. The target decoded layer identifier list, targetDLayerIdList, specifies the list of nuh\_layer\_id values, which need to be decoded for the selected output layer set may be derived as follows:

[0378] For a selected output layer set oLsIdx, a target output layer identifier list targetOpLayerIdList is derived as:

[0379] [Table 12]

for(k=0, numOutputLayers=0; k<=vps_max_layer_id; k++)
if(output_layer_flag[oLsIdx][k])
targetOpLayerIdList [numOutputLayers++] = layer_id_in_nuh[k]

[0380] A target decoded layer identifier list targetDLayerIdList may be derived as:

[0381]

[Table 13]

for(m=0, numDecodedLayers=0; m< numOutputLayers; m++) {
for(n=0; n< NumDirectRefLayers[LayerIdInVps[targetOpLayerIdList[m]]]; n++) {
rLid=RefLayerId[LayerIdInVps[targetOpLayerIdList[m]]][n]
if(rLid not included in targetDLayerIdList[0,..., numDecodedLayers])
targetDLayerIdList[numDecodedLayers++]=rLid;
}
}

- [0382] In some case the final targetDLayerIdList may be obtained by sorting the above targetDLayerIdList in increasing order of nuh\_layer\_id values.
- [0383] A target decoded layer set is identified by the associated target decoded layer identifier list targetDLayerIdList. In some case the targetDLayerIdList may be same as the layer identifier list TargetDecLayerIdList from JCTVC-L1008, which specifies the list of nuh\_layer\_id values, in increasing order of nuh\_layer\_id values, of the NAL units to be decoded:
- [0384] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets if there are no output layer sets change SEI messages.
- [0385] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets + num\_addl\_output\_layer\_sets if there are output layer sets change SEI messages present.
- [0386] Referring to Figure 37, another SEI message syntax may be used to facilitate the change of the output layer sets, namely, output\_layer\_sets\_change(payloadSize).
- [0387] The 'num\_deleted\_output\_layer\_sets' specifies the number of output layer sets which are deleted and so are no longer present. The value of num\_deleted\_output\_layer\_sets is in the range of 0 to num\_output\_layer\_sets, inclusive. When not present, the value of num\_deleted\_output\_layer\_sets is inferred to be equal to 0.
- [0388] The 'deleted\_output\_layer\_sets\_idx\_entry'[ i ] identifies the index entry in the list of output layer set entries which is indicated to be deleted and so no longer present. The value of deleted\_output\_layer\_set\_idx\_entry[ i ] should be in the range of 0 to num\_output\_layer\_sets, inclusive. The deleted\_output\_layer\_sets\_idx\_entry[ i ] is not equal to deleted\_output\_layer\_sets\_idx\_entry[ j ] for any j not equal to i.
- [0389] The 'num\_changed\_output\_layer\_sets' specifies the number of changed output layer sets for which output layers are specified in vps extension section of the active video parameter set identified by active\_vps\_id. The value of

num\_changed\_output\_layer\_sets should be in the range of 0 to num\_output\_layer\_sets - num\_deleted\_output\_layer\_sets, inclusive. When not present, the value of num\_changed\_output\_layer\_sets is inferred to be equal to 0.

[0390] A target list of output layers may be derived, if desired. The target output layer identifier list, targetOpLayerIdList, specifies the list of nuh\_layer\_id values, in increasing order of nuh\_layer\_id values which are output layers for the selected output layer set may be derived as follows:

[0391] For a selected output layer set oLsIdx, a target output layer identifier list targetOpLayerIdList is derived as:

[0392] [Table 14]

for(k=0, numOutputLayers=0; k<=vps_max_layer_id; k++)
if(output_layer_flag[oLsIdx][k])
targetOpLayerIdList [numOutputLayers++] = layer_id_in_nuh[k]

[0393] A target output layer set is identified by the associated target output layer identifier list targetOpLayerIdList.

[0394] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets if there are no output layer sets change SEI messages.

[0395] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets - num\_deleted\_output\_layer\_sets + num\_addl\_output\_layer\_sets if there are output layer sets change SEI messages present.

[0396] A target list of decoded layers may be derived, if desired.

[0397] A target decoded layer identifier list, TargetDLayerIdList, specifies the list of nuh\_layer\_id values, which need to be decoded for the selected output layer set may be derived as follows:

[0398] For a selected output layer set oLsIdx, a target output layer identifier list targetOpLayerIdList may be derived as:

[0399] [Table 15]

for(k=0, numOutputLayers=0; k<=vps_max_layer_id; k++)
if(output_layer_flag[oLsIdx][k])
targetOpLayerIdList [numOutputLayers++] = layer_id_in_nuh[k]

[0400] Then a target decoded layer identifier list targetDLayerIdList may be derived as:

[0401]

[Table 16]

for(m=0, numDecodedLayers=0; m < numOutputLayers; m++) {
for(n=0; n < NumDirectRefLayers[LayerIdInVps[targetOpLayerIdList[m]]]; n++) {
rLid = RefLayerId[LayerIdInVps[targetOpLayerIdList[m]]][n]
if(rLid not included in targetDLayerIdList[0,..., numDecodedLayers])
targetDLayerIdList[numDecodedLayers++] = rLid;
}
}

- [0402] In some case the final targetDLayerIdList may be obtained by sorting the above targetDLayerIdList in increasing order of nuh\_layer\_id values.
- [0403] A target decoded layer set is identified by the associated target decoded layer identifier list targetDLayerIdList. In some case the targetDLayerIdList may be same as the layer identifier list TargetDecLayerIdList from JCTVC-L1008, which specifies the list of nuh\_layer\_id values, in increasing order of nuh\_layer\_id values, of the NAL units to be decoded:
- [0404] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets if there are no output layer sets change SEI messages.
- [0405] The selected output layer set oLsIdx is between 0 and num\_output\_layer\_sets - num\_deleted\_output\_layer\_sets + num\_addl\_output\_layer\_sets if there are output layer sets change SEI messages present.
- [0406] The output layer sets change SEI message is defined such that its effect is not cumulative. In another embodiment this message syntax and/or semantics can be defined such that they are cumulative. In this case the decoder (and encoder) can keep track of the all the previous values of num\_output\_layer\_sets (from vps extension) and/ or num\_deleted\_output\_layer\_sets and/ or num\_changed\_output\_layer\_sets and/ or num\_addl\_output\_layer\_sets and can accumulate these changes every time a new SEI message is signaled.
- [0407] In another embodiment one or more of the syntax elements may be signaled using a known fixed number of bits instead of u(v) instead of ue(v). For example they could be signaled using u(8) or u(16) or u(32) or u(64), etc..
- [0408] In another embodiment one or more of these syntax element could be signaled with ue(v) or some other coding scheme instead of fixed number of bits such as u(v) coding.
- [0409] In another embodiment the names of various syntax elements and their semantics may be altered by adding a plus1 or plus2 or by subtracting a minus1 or a minus2

compared to the described syntax and semantics.

- [0410] In yet another embodiment various syntax elements included in the output layer sets SEI message may be signaled per picture or at other frequency anywhere in the bitstream. For example they may be signaled in slice segment header, pps/ sps/ vps/ adaptation parameter set or any other parameter set or other normative part of the bitstream.
- [0411] In yet another embodiment various syntax elements may be signaled per picture or at other frequency anywhere in the bitstream. For example they may be signaled in slice segment header, pps/ sps/ vps/ adaptation parameter set or any other parameter set or other normative part of the bitstream.
- [0412] In yet another embodiments all the concepts defined in this invention related to output layer sets could be applied to output operation points as defined in JCTVC-L0452 and JCTVC-L0453 and/ or to operation points as defined in JCTVC-L1003.

## **Example 2**

- [0413] Another example of the present invention is described below. Note that, for convenience, members having functions identical to those of the respective members illustrated in the First Example are given respective identical reference numerals, and a description of those members is omitted here.
- [0414] It was determined that signaling the output order decoded picture buffer (DPB) based on using `sps_max_num_reorder_pics[HighestTid]`, `sps_max_latency_increase_plus1[HighestTid]` and `sps_max_dec_pic_buffering[HighestTid]` syntax elements for the output and removal of pictures from the DPB does not account for the buffer characteristics that may result from scalable video coding, such as when different numbers of enhancement layers are used which tends to vary after the content has been encoded based upon the user's viewing preferences, and the multi-view enhancement layers which tends to vary after the content has been encoded based upon the user's viewing preferences. Also it was determined that signaling the output order decoded picture buffer (DPB) based on using `sps_max_num_reorder_pics[HighestTid]`, `sps_max_latency_increase_plus1[HighestTid]` and `sps_max_dec_pic_buffering[HighestTid]` syntax elements for the output and removal of pictures from the DPB may not be optimal in terms of the memory usage of the DPB when decoder operates at a certain operation point and/ or is outputting selected output layer set. To accommodate such differences in the viewing preferences, the output order decoded picture buffer (DPB) may further and/or alternatively be based upon such syntax elements being included together with the video parameter set extension (VPS extension) to provide syntax elements for one or more of the en-

hancement layers. In this manner the syntax elements may be selected to be especially suitable for the particular operation point or output layer set, which tends to correspond to the user's viewing preferences.

- [0415] The DPB buffering related parameters, `vps_max_dec_pic_buffering_minus1`, `vps_max_num_reorder_pics`, `vps_max_latency_increase_plus1` may be signaled for sub-layers for the CVS for one or more operation points and/or for output layer sets in VPS extension. Similarly, the system may define the operation and bumping process for the output order DPB to use the above signalled DPB buffering parameters from the VPS extension if they are signalled for the operation point under test or for the selected output layer set. Otherwise the corresponding SPS level parameters from the active SPS (when `currLayerId` which corresponds to `nuh_layer_id` of the current picture is equal to 0) or from the active layer SPS depending upon the `layer_id` of the current layer are used. In some cases the parameters `vps_max_dec_pic_buffering_minus1`, `vps_max_num_reorder_pics`, `vps_max_latency_increase_plus1` may be termed as `max_vps_dec_pic_buffering_minus1`, `max_vps_num_reorder_pics`, `max_vps_latency_increase_plus1`.
- [0416] In another case the DPB buffering related parameter `max_vps_dec_pic_buffering_minus1` is signaled for sub-layers for the CVS for one or more output layer sets in VPS extension. The output layer sets may correspond to operating points. The DPB buffering related parameters `max_vps_num_reorder_pics`, `max_vps_latency_increase_plus1` may be signaled for sub-layers for the CVS for each layer in VPS extension. Similarly, the system may define the operation and bumping process for the output order DPB to use the above signalled DPB buffering parameters from the VPS extension if they are signalled for the operation point under test or for the selected output layer set. Otherwise the corresponding SPS level parameters from the active SPS (when `currLayerId` which corresponds to `nuh_layer_id` of the current picture is equal to 0) or from the active layer SPS depending upon the `layer_id` of the current layer are used.
- [0417] Referring to Figure 38A, an exemplary modified `vps_extension` is illustrated. The modified `vps` extension includes new syntax, namely, `num_op_dpb_info_parameters` and `operation_point_layer_set_idx[i]`. This modified `vps` extension may be defined in terms of the operation point which is a bitstream created from another bitstream by operation of a sub-bitstream extraction process with the another bitstream, a target highest `TemporalId`, and a target layer identifier list as inputs.
- [0418] `num_output_layer_sets` specifies the number of layer sets for which output layers are specified with `output_layer_set_index[i]` and `output_layer_flag[lsIdx][j]`. When not present, the value of `num_output_layer_sets` is inferred to be equal to 0. A layer set de-

scribing output layers is an output layer set.

- [0419] `output_layer_set_idx[ i ]` specifies the index `lsIdx` of the layer set for which `output_layer_flag[ lsIdx ][ j ]` is present.
- [0420] `output_layer_flag[ lsIdx ][ j ]` equal to 1 specifies that the layer with `nuh_layer_id` equal to `j` is a target output layer of the `lsIdx`-th layer set. A value of `output_layer_flag[ lsIdx ][ j ]` equal to 0 specifies that the layer with `nuh_layer_id` equal to `j` is not a target output layer of the `lsIdx`-th layer set.
- [0421] The `num_op_dpb_info_parameters` specifies the number of `op_dpb_parameters( )` syntax structures present in the VPS extension RBSP, defined in terms of the operation point. The `num_op_dpb_info_parameters` decoders is in the range of 0 to `vps_num_layer_sets_minus1`, inclusive.
- [0422] The `operation_point_layer_set_idx[ i ]` specifies the index, into the list of layer sets defined by operation points to which the `i` th `op_dpb_info_parameters( )` syntax structure in the VPS extension applies. The value of `operation_point_layer_set_idx[ i ]` may be in the range of 0 to `vps_num_layer_sets_minus1`, inclusive. For bitstream conformance the `operation_point_layer_set_idx[ i ]` is not equal to `operation_point_layer_set_idx[ j ]` for any `j` not equal to `i`.
- [0423] Referring to Figure 39A, the `op_dpb_info_parameters` specifies `vps_max_sub_layers_minus1[ j ]`, `vps_sub_layer_ordering_info_present_flag[ j ]`, `vps_max_dec_pic_buffering_minus1[ j ][ k ]`, `vps_max_num_reorder_pics[ j ][ k ]`, and `vps_max_latency_increase_plus1[ j ][ k ]`.
- [0424] The `vps_max_sub_layers_minus1[ j ]` plus 1 indicates how many sub layers are included. The `vps_max_sub_layers_minus1[ j ]` plus 1 specifies the maximum number of temporal sub-layers that may be present in the CVS for layer with `nuh_layer_id` equal to `j`. The value of `vps_max_sub_layers_minus1[ j ]` is in the range of 0 to 6, inclusive.
- [0425] The `vps_sub_layer_ordering_info_present_flag[ j ]` indicates whether the syntax is for one set including all layers or for each individual layer. The `vps_sub_layer_ordering_info_present_flag[ j ]` equal to 1 specifies that `vps_max_dec_pic_buffering_minus1[ j ][ k ]`, `vps_max_num_reorder_pics[ j ][ k ]`, and `vps_max_latency_increase_plus1[ j ][ k ]` are present for layer with `nuh_layer_id` equal to `j` for `vps_max_sub_layers_minus1[ j ] + 1` sub-layers. The `vps_sub_layer_ordering_info_present_flag[ j ]` equal to 0 specifies that the values of `vps_max_dec_pic_buffering_minus1[ j ][ vps_max_sub_layers_minus1[ j ] ]`, `vps_max_num_reorder_pics[ j ][ vps_max_sub_layers_minus1[ j ] ]`, and `vps_max_latency_increase_plus1[ j ][ vps_max_sub_layers_minus1[ j ] ]` apply to all sub-layers for layer with `nuh_layer_id` equal to `j`.
- [0426] The `vps_max_dec_pic_buffering_minus1[ j ][ k ]` plus 1 specifies the maximum



required size of the decoded picture buffer for the CVS for layer with nuh\_layer\_id equal to j in units of picture storage buffers when HighestTid is equal to k. The value of vps\_max\_dec\_pic\_buffering\_minus1[j][k] shall be in the range of 0 to MaxDpbSize - 1 (as specified in subclause A.4), inclusive. When k is greater than 0, vps\_max\_dec\_pic\_buffering\_minus1[j][k] shall be greater than or equal to vps\_max\_dec\_pic\_buffering\_minus1[j][k - 1]. When vps\_max\_dec\_pic\_buffering\_minus1[j][k] is not present for k in the range of 0 to vps\_max\_sub\_layers\_minus1[j] - 1, inclusive, due to vps\_sub\_layer\_ordering\_info\_present\_flag[j] being equal to 0, it is inferred to be equal to vps\_max\_dec\_pic\_buffering\_minus1[j][vps\_max\_sub\_layers\_minus1[j]].

[0427] The vps\_max\_num\_reorder\_pics[j][k] indicates the maximum allowed number of pictures that can precede any picture in the CVS for layer with nuh\_layer\_id equal to j in decoding order and follow that picture in output order when HighestTid is equal to k. The value of vps\_max\_num\_reorder\_pics[j][k] shall be in the range of 0 to vps\_max\_dec\_pic\_buffering\_minus1[j][k], inclusive. When k is greater than 0, vps\_max\_num\_reorder\_pics[j][k] is greater than or equal to vps\_max\_num\_reorder\_pics[j][k - 1]. When vps\_max\_num\_reorder\_pics[j][k] is not present for k in the range of 0 to vps\_max\_sub\_layers\_minus1[j] - 1, inclusive, due to vps\_sub\_layer\_ordering\_info\_present\_flag[j] being equal to 0, it is inferred to be equal to vps\_max\_num\_reorder\_pics[j][vps\_max\_sub\_layers\_minus1[j]].

[0428] The vps\_max\_latency\_increase\_plus1[j][k] not equal to 0 is used to compute the value of VpsMaxLatencyPictures[j][k], which specifies the maximum number of pictures that can precede any picture in the CVS for layer with nuh\_layer\_id equal to j in output order and follow that picture in decoding order when HighestTid is equal to k.

[0429] When vps\_max\_latency\_increase\_plus1[j][k] is not equal to 0, the value of VpsMaxLatencyPictures[j][k] may be specified as follows:

[0430] 
$$\text{VpsMaxLatencyPictures}[j][k] = \text{vps\_max\_num\_reorder\_pics}[j][k] + \text{vps\_max\_latency\_increase\_plus1}[j][k] - 1$$

[0431] When vps\_max\_latency\_increase\_plus1[j][k] is equal to 0, no corresponding limit is expressed.

[0432] The value of vps\_max\_latency\_increase\_plus1[j][k] is in the range of 0 to  $2^{32} - 2$ , inclusive. When vps\_max\_latency\_increase\_plus1[j][k] is not present for k in the range of 0 to vps\_max\_sub\_layers\_minus1[j] - 1, inclusive, due to vps\_sub\_layer\_ordering\_info\_present\_flag[j] being equal to 0, it is inferred to be equal to vps\_max\_latency\_increase\_plus1[j][vps\_max\_sub\_layers\_minus1[j]].

[0433] The 'vps\_max\_sub\_layers\_minus1'[id][j] plus 1 specifies the maximum number of temporal sub-layers that may be present in the CVS for layer with nuh\_layer\_id equal

- to  $j$  for the operation point associated with index  $id$ . The value of  $vps\_max\_sub\_layers\_minus1[id][j]$  shall be in the range of 0 to 6, inclusive.
- [0434] The ' $vps\_sub\_layer\_ordering\_info\_present\_flag[id][j]$ ' equal to 1 specifies that  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k]$ ,  $vps\_max\_num\_reorder\_pics[id][j][k]$ , and  $vps\_max\_latency\_increase\_plus1[id][j][k]$  are present for layer with  $nuh\_layer\_id$  equal to  $j$  for the operation point associated with index  $id$  for  $vps\_max\_sub\_layers\_minus1[id][j] + 1$  sub-layers.
- [0435]  $vps\_sub\_layer\_ordering\_info\_present\_flag[id][j]$  equal to 0 specifies that the values of  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][vps\_max\_sub\_layers\_minus1[id][j]]$ ,  $vps\_max\_num\_reorder\_pics[id][j][vps\_max\_sub\_layers\_minus1[id][j]]$ , and  $vps\_max\_latency\_increase\_plus1[id][j][vps\_max\_sub\_layers\_minus1[id][j]]$  apply to all sub-layers for layer with  $nuh\_layer\_id$  equal to  $j$  for the operation point associated with index  $id$ .
- [0436] The ' $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k]$ ' plus 1 specifies the maximum required size of the decoded picture buffer for the CVS for layer with  $nuh\_layer\_id$  equal to  $j$  for the operation point associated with index  $id$  in units of picture storage buffers when  $HighestTid$  is equal to  $k$ . The value of  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k]$  shall be in the range of 0 to  $MaxDpbSize - 1$  (as specified in subclause A.4), inclusive. When  $k$  is greater than 0,  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k]$  shall be greater than or equal to  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k - 1]$ . When  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k]$  is not present for  $k$  in the range of 0 to  $vps\_max\_sub\_layers\_minus1[id][j] - 1$ , inclusive, due to  $vps\_sub\_layer\_ordering\_info\_present\_flag[id][j]$  being equal to 0, it is inferred to be equal to  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][vps\_max\_sub\_layers\_minus1[id][j]]$ .
- [0437] The ' $vps\_max\_num\_reorder\_pics[id][j][k]$ ' indicates the maximum allowed number of pictures that can precede any picture in the CVS for layer with  $nuh\_layer\_id$  equal to  $j$  for the operation point associated with index  $id$  in decoding order and follow that picture in output order when  $HighestTid$  is equal to  $k$ . The value of  $vps\_max\_num\_reorder\_pics[id][j][k]$  shall be in the range of 0 to  $vps\_max\_dec\_pic\_buffering\_minus1[id][j][k]$ , inclusive. When  $k$  is greater than 0,  $vps\_max\_num\_reorder\_pics[id][j][k]$  shall be greater than or equal to  $vps\_max\_num\_reorder\_pics[id][j][k - 1]$ . When  $vps\_max\_num\_reorder\_pics[id][j][k]$  is not present for  $k$  in the range of 0 to  $vps\_max\_sub\_layers\_minus1[id][j] - 1$ , inclusive, due to  $vps\_sub\_layer\_ordering\_info\_present\_flag[id][j]$  being equal to 0, it is inferred to be equal to  $vps\_max\_num\_reorder\_pics[id][j][vps\_max\_sub\_layers\_minus1[id][j]]$ .

- [0438] The 'vps\_max\_latency\_increase\_plus1'[ id ][ j ][ k ] not equal to 0 is used to compute the value of VpsMaxLatencyPictures[ id ][ j ][ k ], which specifies the maximum number of pictures that can precede any picture in the CVS for layer with nuh\_layer\_id equal to j for the operation point associated with index id in output order and follow that picture in decoding order when HighestTid is equal to k.
- [0439] When vps\_max\_latency\_increase\_plus1[ id ][ j ][ k ] is not equal to 0, the value of VpsMaxLatencyPictures[ id ][ j ][ k ] is specified as follows:
- [0440] 
$$\text{VpsMaxLatencyPictures}[ \text{id} ][ \text{j} ][ \text{k} ] = \text{vps\_max\_num\_reorder\_pics}[ \text{id} ][ \text{j} ][ \text{k} ] + \text{vps\_max\_latency\_increase\_plus1}[ \text{id} ][ \text{j} ][ \text{k} ] - 1$$
- [0441] When vps\_max\_latency\_increase\_plus1[ id ][ j ][ k ] is equal to 0, no corresponding limit is expressed.
- [0442] The value of vps\_max\_latency\_increase\_plus1[ id ][ j ][ k ] shall be in the range of 0 to  $2^{32} - 2$ , inclusive. When vps\_max\_latency\_increase\_plus1[ id ][ j ][ k ] is not present for k in the range of 0 to vps\_max\_sub\_layers\_minus1[ id ][ j ] - 1, inclusive, due to vps\_sub\_layer\_ordering\_info\_present\_flag[ id ][ j ] being equal to 0, it is inferred to be equal to vps\_max\_latency\_increase\_plus1[ id ][ j ][ vps\_max\_sub\_layers\_minus1[ id ][ j ] ].
- [0443] Referring to Figure 39 B, the op\_dpb\_info\_parameters may be further modified as shown to op\_dpb\_info\_parameters(id,j). In this case the syntax of VPS extension may be as illustrated in Figure 38B. The hypothetical reference decoder (HRD) is used to check bitstream and decoder conformance. Two types of bitstreams or bitstream subsets are subject to HRD conformance checking for the Joint Collaborative Team on Video Coding (JCT-VC). The first type, called a Type I bitstream, is a NAL unit stream containing only the VCL NAL units and NAL units with nal\_unit\_type equal to FD\_NUT (filler data NAL units) for all access units in the bitstream. The second type, called a Type II bitstream, contains, in addition to the VCL NAL units and filler data NAL units for all access units in the bitstream, at least one of (a) additional non-VCL NAL units other than filler data NAL units, and (b) all leading\_zero\_8bits, zero\_byte, start\_code\_prefix\_one\_3bytes, and trailing\_zero\_8bits syntax elements that form a byte stream from the NAL unit stream.
- [0444] The syntax elements of non-VCL NAL units (or their default values for some of the syntax elements), required for the HRD, are specified in the semantic subclauses of clause 7, Annexes D and E.
- [0445] Two types of HRD parameter sets (NAL HRD parameters and VCL HRD parameters) are used. The HRD parameter sets are signalled through the hrd\_parameters( ) syntax structure, which may be part of the SPS syntax structure or the VPS syntax structure.
- [0446] Multiple tests may be needed for checking the conformance of a bitstream, which is

referred to as the bitstream under test. For each test, the following steps apply in the order listed:

- [0447] (1) An operation point under test, denoted as TargetOp, is selected. The layer identifier list OpLayerIdList of TargetOp consists of the list of nuh\_layer\_id values, in increasing order of nuh\_layer\_id values, present in the bitstream subset associated with TargetOp, which is a subset of the nuh\_layer\_id values present in the bitstream under test. The OpTid of TargetOp is equal to the highest TemporalId present in the bitstream subset associated with TargetOp.
- [0448] (2) TargetDecLayerIdList is set equal to OpLayerIdList of TargetOp, HighestTid is set equal to OpTid of TargetOp, and the sub-bitstream extraction process as specified in clause 10 is invoked with the bitstream under test, HighestTid, and TargetDecLayerIdList as inputs, and the output is assigned to BitstreamToDecode.
- [0449] (3) The hrd\_parameters( ) syntax structure and the sub\_layer\_hrd\_parameters( ) syntax structure applicable to TargetOp are selected. If TargetDecLayerIdList contains all nuh\_layer\_id values present in the bitstream under test, the hrd\_parameters( ) syntax structure in the active SPS (or provided through an external means not specified in this Specification) is selected. Otherwise, the hrd\_parameters( ) syntax structure in the active VPS (or provided through some external means not specified in this Specification) that applies to TargetOp is selected. Within the selected hrd\_parameters( ) syntax structure, if BitstreamToDecode is a Type I bitstream, the sub\_layer\_hrd\_parameters( HighestTid ) syntax structure that immediately follows the condition "if( vcl\_hrd\_parameters\_present\_flag )" is selected and the variable NalHrdModeFlag is set equal to 0; otherwise (BitstreamToDecode is a Type II bitstream), the sub\_layer\_hrd\_parameters( HighestTid ) syntax structure that immediately follows either the condition "if( vcl\_hrd\_parameters\_present\_flag )" (in this case the variable NalHrdModeFlag is set equal to 0) or the condition "if( nal\_hrd\_parameters\_present\_flag )" (in this case the variable NalHrdModeFlag is set equal to 1) is selected. When BitstreamToDecode is a Type II bitstream and NalHrdModeFlag is equal to 0, all non-VCL NAL units except filler data NAL units, and all leading\_zero\_8bits, zero\_byte, start\_code\_prefix\_one\_3bytes, and trailing\_zero\_8bits syntax elements that form a byte stream from the NAL unit stream (as specified in Annex B), when present, are discarded from BitstreamToDecode, and the remaining bitstream is assigned to BitstreamToDecode.
- [0450] In another case Multiple tests may be needed for checking the conformance of a bitstream, which is referred to as the bitstream under test. For each test, the following steps apply in the order listed:
- [0451] (1) An output layer set under test, denoted as TargetOpLs is selected. The operation point referred in TargetOpLs by output\_layer\_set\_idx[] identifies the operation point

under test. The output layer identifier list `OpLayerIdList` of `TargetOpLs` consists of the list of `nuh_layer_id` values, in increasing order of `nuh_layer_id` values, present in the bitstream subset associated with `TargetOp` and `TargetOpLs`, which is a subset of the `nuh_layer_id` values present in the bitstream under test. The `OpTid` of `TargetOp` is equal to the highest `TemporalId` present in the bitstream subset associated with `TargetOp`.

- [0452] (2) `TargetDecLayerIdList` is set equal to target decoded layer identifier list `targetDecLayerIdList` for the selected output layer set `TargetOpLs`, `HighestTid` is set equal to `OpTid` of `TargetOp`, and the sub-bitstream extraction process as specified in clause 10 is invoked with the bitstream under test, `HighestTid`, and `TargetDecLayerIdList` as inputs, and the output is assigned to `BitstreamToDecode`.
- [0453] (3) The `hrd_parameters()` syntax structure and the `sub_layer_hrd_parameters()` syntax structure applicable to `TargetOp` are selected. If `TargetDecLayerIdList` contains all `nuh_layer_id` values present in the bitstream under test, the `hrd_parameters()` syntax structure in the active SPS (or provided through an external means not specified in this Specification) is selected. Otherwise, the `hrd_parameters()` syntax structure in the active VPS (or provided through some external means not specified in this Specification) that applies to `TargetOp` is selected. Within the selected `hrd_parameters()` syntax structure, if `BitstreamToDecode` is a Type I bitstream, the `sub_layer_hrd_parameters(HighestTid)` syntax structure that immediately follows the condition "`if( vcl_hrd_parameters_present_flag )`" is selected and the variable `NalHrdModeFlag` is set equal to 0; otherwise (`BitstreamToDecode` is a Type II bitstream), the `sub_layer_hrd_parameters(HighestTid)` syntax structure that immediately follows either the condition "`if( vcl_hrd_parameters_present_flag )`" (in this case the variable `NalHrdModeFlag` is set equal to 0) or the condition "`if( nal_hrd_parameters_present_flag )`" (in this case the variable `NalHrdModeFlag` is set equal to 1) is selected. When `BitstreamToDecode` is a Type II bitstream and `NalHrdModeFlag` is equal to 0, all non-VCL NAL units except filler data NAL units, and all `leading_zero_8bits`, `zero_byte`, `start_code_prefix_one_3bytes`, and `trailing_zero_8bits` syntax elements that form a byte stream from the NAL unit stream (as specified in Annex B), when present, are discarded from `BitstreamToDecode`, and the remaining bitstream is assigned to `BitstreamToDecode`.
- [0454] A conforming decoder may fulfil all requirements specified in this subclause.
- [0455] (1) A decoder claiming conformance to a specific profile, tier and level shall be able to successfully decode all bitstreams that conform to the bitstream conformance requirements specified in subclause C.4, in the manner specified in Annex A, provided that all VPSs, SPSs and PPSs referred to in the VCL NAL units, and appropriate buffering period and picture timing SEI messages are conveyed to the decoder, in a

timely manner, either in the bitstream (by non-VCL NAL units), or by external means not specified in this Specification.

- [0456] (2) When a bitstream contains syntax elements that have values that are specified as reserved and it is specified that decoders shall ignore values of the syntax elements or NAL units containing the syntax elements having the reserved values, and the bitstream is otherwise conforming to this Specification, a conforming decoder shall decode the bitstream in the same manner as it would decode a conforming bitstream and shall ignore the syntax elements or the NAL units containing the syntax elements having the reserved values as specified.
- [0457] There are two types of conformance of a decoder: output timing conformance and output order conformance.
- [0458] To check conformance of a decoder, test bitstreams conforming to the claimed profile, tier and level, as specified in subclause C.4 are delivered by a hypothetical stream scheduler (HSS) both to the HRD and to the decoder under test (DUT). All cropped decoded pictures output by the HRD shall also be output by the DUT, each cropped decoded picture output by the DUT shall be a picture with PicOutputFlag equal to 1, and, for each such cropped decoded picture output by the DUT, the values of all samples that are output shall be equal to the values of the samples produced by the specified decoding process.
- [0459] For output timing decoder conformance, the HSS operates as described above, with delivery schedules selected only from the subset of values of SchedSelIdx for which the bit rate and CPB size are restricted as specified in Annex A for the specified profile, tier and level, or with "interpolated" delivery schedules as specified below for which the bit rate and CPB size are restricted as specified in Annex A. The same delivery schedule is used for both the HRD and the DUT.
- [0460] When the HRD parameters and the buffering period SEI messages are present with  $\text{cpb\_cnt\_minus1}[\text{HighestTid}]$  greater than 0, the decoder shall be capable of decoding the bitstream as delivered from the HSS operating using an "interpolated" delivery schedule specified as having peak bit rate  $r$ , CPB size  $c(r)$ , and initial CPB removal delay
- [0461] [Math.15]
- $$(f(r) \div r)$$
- as follows:

[Table 17]

$$\alpha = (r - \text{BitRate}[\text{SchedSelIdx} - 1]) \div (\text{BitRate}[\text{SchedSelIdx}] - \text{BitRate}[\text{SchedSelIdx} - 1]), \quad (\text{C-22})$$

$$c(r) = \alpha * \text{CpbSize}[\text{SchedSelIdx}] + (1 - \alpha) * \text{CpbSize}[\text{SchedSelIdx} - 1], \quad (\text{C-23})$$

$$f(r) = \alpha * \text{InitCpbRemovalDelay}[\text{SchedSelIdx}] * \text{BitRate}[\text{SchedSelIdx}] + (1 - \alpha) * \text{InitCpbRemovalDelay}[\text{SchedSelIdx} - 1] * \text{BitRate}[\text{SchedSelIdx} - 1] \quad (\text{C-24})$$

- [0462] for any SchedSelIdx > 0 and r such that BitRate[ SchedSelIdx - 1 ] <= r <= BitRate[ SchedSelIdx ] such that r and c( r ) are within the limits as specified in Annex A for the maximum bit rate and buffer size for the specified profile, tier and level. The InitCpbRemovalDelay[ SchedSelIdx ] can be different from one buffering period to another and have to be re-calculated.
- [0463] For output timing decoder conformance, an HRD as described above is used and the timing (relative to the delivery time of the first bit) of picture output is the same for both the HRD and the DUT up to a fixed delay.
- [0464] For output order decoder conformance, the following applies:
- [0465] (1) The HSS delivers the bitstream BitstreamToDecode to the DUT "by demand" from the DUT, meaning that the HSS delivers bits (in decoding order) only when the DUT requires more bits to proceed with its processing. This means that for this test, the coded picture buffer of the DUT could be as small as the size of the largest decoding unit.
- [0466] (2) A modified HRD as described below is used, and the HSS delivers the bitstream to the HRD by one of the schedules specified in the bitstream BitstreamToDecode such that the bit rate and CPB size are restricted as specified in Annex A. The order of pictures output shall be the same for both the HRD and the DUT.
- [0467] (3) The HRD CPB size is given by CpbSize[ SchedSelIdx ] as specified in subclause E.2.3, where SchedSelIdx and the HRD parameters are selected as specified in subclause C.1. The DPB size is given by sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] + 1 from the active SPS (when nuh\_layer\_id for the current decoded picture is equal to 0) or from the active layer SPS for the value of nuh\_layer\_id of the current decoded picture. In some cases, if operation point DPB information parameters op\_dpb\_info\_parameters() are present for the selected output layer set, The DPB size is given by vps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] when currLayerId is equal to 0 or is set to vps\_max\_dec\_pic\_buffering\_minus1[ CurrLayerId ][ HighestTid ] for the currLayerId for the operation point under test when currLayerId is greater than 0, where currLayerId is the nuh\_layer\_id of the current decoded picture. Otherwise if operation point DPB information parameters op\_dpb\_info\_parameters() are not present for the operation point under test, the DPB Size is given by sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] + 1 from the active SPS (when

nuh\_layer\_id for the current decoded picture is equal to 0) or from the active layer SPS for the value of nuh\_layer\_id of the current decoded picture.

- [0468] In some cases, if output layer sets DPB information parameters oop\_dpb\_info\_parameters() are present for the selected output layer set, The DPB size is given by vps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] when currLayerId is equal to 0 or is set to vps\_max\_dec\_pic\_buffering\_minus1[ CurrLayerId ][ HighestTid ] for the currLayerId for the selected output layer set, where currLayerId is the nuh\_layer\_id of the current decoded picture. Otherwise if output layer sets DPB information parameters oop\_dpb\_info\_parameters() are not present for the selected output layer set, the DPB Size is given by sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] + 1 from the active SPS (when nuh\_layer\_id for the current decoded picture is equal to 0) or from the active layer SPS for the value of nuh\_layer\_id of the current decoded picture.
- [0469] The removal time from the CPB for the HRD is the final bit arrival time and decoding is immediate. The operation of the DPB of this HRD is as described in subclauses C.5.2 through C.5.2.3.
- [0470] The decoded picture buffer contains picture storage buffers. The number of picture storage buffers for nuh\_layer\_id equal to 0 is derived from the active SPS. The number of picture storage buffers for each non-zero nuh\_layer\_id value is derived from the active layer SPS for that non-zero nuh\_layer\_id value. Each of the picture storage buffers contains a decoded picture that is marked as "used for reference" or is held for future output. The process for output and removal of pictures from the DPB as specified in subclause F.13.5.2.2 is invoked, followed by the invocation of the process for picture decoding, marking, additional bumping, and storage as specified in subclause F.13.5.2.3. The "bumping" process is specified in subclause F.13.5.2.4 and is invoked as specified in subclauses F.13.5.2.2 and F.13.5.2.3.
- [0471] The output and removal of pictures from the DPB before the decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously when the first decoding unit of the access unit containing the current picture is removed from the CPB and proceeds as follows.
- [0472] The decoding process for RPS as specified in subclause 8.3.2 is invoked.
- [0473] (1) If the current picture is an IRAP picture with NoRaslOutputFlag equal to 1 and with nuh\_layer\_id equal to 0 that is not picture 0, the following ordered steps are applied:
- [0474] (A) The variable NoOutputOfPriorPicsFlag is derived for the decoder under test as follows:
- [0475] (i) If the current picture is a CRA picture, NoOutputOfPriorPicsFlag is set equal to 1 (regardless of the value of no\_output\_of\_prior\_pics\_flag).



- [0476] (ii) Otherwise, if the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]` derived from the active SPS is different from the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, respectively, derived from the SPS active for the preceding picture, `NoOutputOfPriorPicsFlag` may (but should not) be set to 1 by the decoder under test, regardless of the value of `no_output_of_prior_pics_flag`. Although setting `NoOutputOfPriorPicsFlag` equal to `no_output_of_prior_pics_flag` is preferred under these conditions, the decoder under test is allowed to set `NoOutputOfPriorPicsFlag` to 1 in this case.
- [0477] (iii) Otherwise, `NoOutputOfPriorPicsFlag` is set equal to `no_output_of_prior_pics_flag`.
- [0478] (B) The value of `NoOutputOfPriorPicsFlag` derived for the decoder under test is applied for the HRD as follows:
- [0479] (i) If `NoOutputOfPriorPicsFlag` is equal to 1, all picture storage buffers in the DPB are emptied without output of the pictures they contain, and the DPB fullness is set equal to 0.
- [0480] (ii) Otherwise (`NoOutputOfPriorPicsFlag` is equal to 0), all picture storage buffers containing a picture that is marked as "not needed for output" and "unused for reference" are emptied (without output), and all non-empty picture storage buffers in the DPB are emptied by repeatedly invoking the "bumping" process specified in subclause F.13.5.2.4, and the DPB fullness is set equal to 0.
- [0481] (iii) Otherwise (the current picture is not an IRAP picture with `NoRaslOutputFlag` equal to 1 and with `nuh_layer_id` equal to 0), all picture storage buffers containing a picture which are marked as "not needed for output" and "unused for reference" are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness is decremented by one. The variable `currLayerId` is set equal to `nuh_layer_id` of the current decoded picture.
- [0482] The variables `MaxNumReorderPics[ TargetOp ][ currLayerId ][ HighestTid ]`, `MaxLatencyIncreasePlus1[ TargetOp ][ currLayerId ][ HighestTid ]`, `MaxLatencyPictures[ TargetOp ][ currLayerId ][ HighestTid ]`, `MaxDecPicBufferingMinus1[ TargetOp ][ currLayerId ][ HighestTid ]` are derived as follows based on the current operation point under test:
- [0483] (1) If operation point DPB information parameters `op_dpb_info_parameters()` are present for the operation point under test `TargetOp`, `MaxNumReorderPics[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `vps_max_num_reorder_pics[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `vps_max_num_reorder_pics[ TargetOp ][ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. Otherwise if operation point DPB information pa-

rameters `op_dpb_info_parameters()` are not present for the operation point under test `MaxNumReorderPics[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `sps_max_num_reorder_pics[ HighestTid ]` from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.

- [0484] (2) If operation point DPB information parameters `op_dpb_info_parameters()` are present for the operation point under test `TargetOp`, `MaxLatencyIncreasePlus1[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `vps_max_latency_increase_plus1[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `vps_max_latency_increase_plus1[ TargetOp ][ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. If operation point DPB information parameters `op_dpb_info_parameters()` are present for the operation point under test, `MaxLatencyPictures[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `VpsMaxLatencyPictures[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `VpsMaxLatencyPictures[ TargetOp ][ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. Otherwise if operation point DPB information parameters `op_dpb_info_parameters()` are not present for the operation point under test, `MaxLatencyIncreasePlus1[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `sps_max_latency_increase_plus1[ HighestTid ]` of the active SPS (when `currLayerId` is equal to 0) or the active layer SPS for the value of `currLayerId` and `MaxLatencyPictures[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `SpsMaxLatencyPictures[ HighestTid ]` derived from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.

- [0485] (3) If operation point DPB information parameters `op_dpb_info_parameters()` are present for the selected operation point under test `TargetOp`, `MaxDecPicBufferingMinus1[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `vps_max_dec_pic_buffering_minus1[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `vps_max_dec_pic_buffering_minus1[ TargetOp ][ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. Otherwise if operation point DPB information parameters `op_dpb_info_parameters()` are not present for the operation point under test, `MaxDecPicBufferingMinus1[ TargetOp ][ currLayerId ][ HighestTid ]` is set to `sps_max_dec_pic_buffering_minus1[ HighestTid ]` from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.

- [0486] When one or more of the following conditions are true, the "bumping" process specified in subclause F.13.5.2.4 is invoked repeatedly while further decrementing the DPB fullness by one for each additional picture storage buffer that is emptied, until none of the following conditions are true:

- [0487] (1) The number of pictures with `nuh_layer_id` equal to `currLayerId` in the DPB that

- are marked as "needed for output" is greater than  $\text{MaxNumReorderPics}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$ .
- [0488] (2) If  $\text{MaxLatencyIncreasePlus1}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$  is not equal to 0 and there is at least one picture with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB that is marked as "needed for output" for which the associated variable  $\text{PicLatencyCount}[\text{currLayerId}]$  is greater than or equal to  $\text{MaxLatencyPictures}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$ .
- [0489] (3) The number of pictures with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB is greater than or equal to  $\text{MaxDecPicBuffering}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$ .
- [0490] The processes specified in this subclause happen instantaneously when the last decoding unit of access unit  $n$  containing the current picture is removed from the CPB.
- [0491] The variable  $\text{currLayerId}$  is set equal to  $\text{nuh\_layer\_id}$  of the current decoded picture.
- [0492] For each picture in the DPB that is marked as "needed for output" and that has a  $\text{nuh\_layer\_id}$  value equal to  $\text{currLayerId}$ , the associated variable  $\text{PicLatencyCount}[\text{currLayerId}]$  is set equal to  $\text{PicLatencyCount}[\text{currLayerId}] + 1$ .
- [0493] The current picture is considered as decoded after the last decoding unit of the picture is decoded. The current decoded picture is stored in an empty picture storage buffer in the DPB, and the following applies:
- [0494] (A) If the current decoded picture has  $\text{PicOutputFlag}$  equal to 1, it is marked as "needed for output" and its associated variable  $\text{PicLatencyCount}[\text{currLayerId}]$  is set equal to 0.
- [0495] (B) Otherwise (the current decoded picture has  $\text{PicOutputFlag}$  equal to 0), it is marked as "not needed for output".
- [0496] The current decoded picture is marked as "used for short-term reference".
- [0497] When one or more of the following conditions are true, the "bumping" process specified in subclause F.13.5.2.4 is invoked repeatedly until none of the following conditions are true.
- [0498] (A) The number of pictures with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB that are marked as "needed for output" is greater than  $\text{MaxNumReorderPics}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$ .
- [0499] (B)  $\text{MaxLatencyIncreasePlus1}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$  is not equal to 0 and there is at least one picture with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB that is marked as "needed for output" for which the associated variable  $\text{PicLatencyCount}[\text{currLayerId}]$  is greater than or equal to  $\text{MaxLatencyPictures}[\text{TargetOp}][\text{CurrLayerId}][\text{HighestTid}]$ .
- [0500] In other case the variables  $\text{MaxNumReorderPics}[\text{currLayerId}][\text{HighestTid}]$ ,  $\text{MaxLatencyIncreasePlus1}[\text{currLayerId}][\text{HighestTid}]$ ,  $\text{MaxLatencyPictures}[\text{currLayerId}][\text{HighestTid}]$ ,  $\text{MaxDecPicBufferingMinus1}[\text{currLayerId}][\text{HighestTid}]$

may be derived as follows:

- [0501] (1) If operation point DPB information parameters `op_dpb_info_parameters()` are present for the operation point under test, `MaxNumReorderPics[ currLayerId ][ HighestTid ]` is set to `vps_max_num_reorder_pics[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `vps_max_num_reorder_pics[ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. Otherwise if operation point DPB information parameters `op_dpb_info_parameters()` are not present for the operation point under test `MaxNumReorderPics[ currLayerId ][ HighestTid ]` is set to `sps_max_num_reorder_pics[ HighestTid ]` from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.
- [0502] (2) If operation point DPB information parameters `op_dpb_info_parameters()` are present for the operation point under test, `MaxLatencyIncreasePlus1[ currLayerId ][ HighestTid ]` is set to `vps_max_latency_increase_plus1[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `vps_max_latency_increase_plus1[ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. If operation point DPB information parameters `op_dpb_info_parameters()` are present for the operation point under test, `MaxLatencyPictures[ currLayerId ][ HighestTid ]` is set to `VpsMaxLatencyPictures [ HighestTid ]` when `currLayerId` is equal to 0 or is set to `VpsMaxLatencyPictures [ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. Otherwise if operation point DPB information parameters `op_dpb_info_parameters()` are not present for the operation point under test, `MaxLatencyIncreasePlus1[ currLayerId ][ HighestTid ]` is set to `sps_max_latency_increase_plus1[ HighestTid ]` of the active SPS (when `currLayerId` is equal to 0) or the active layer SPS for the value of `currLayerId` and `MaxLatencyPictures[ currLayerId ][ HighestTid ]` is set to `SpsMaxLatencyPictures [ HighestTid ]` derived from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.
- [0503] (3) If operation point DPB information parameters `op_dpb_info_parameters()` are present for the selected operation point under test, `MaxDecPicBufferingMinus1[ currLayerId ][ HighestTid ]` is set to `vps_max_dec_pic_buffering_minus1[ HighestTid ]` when `currLayerId` is equal to 0 or is set to `vps_max_dec_pic_buffering_minus1[ CurrLayerId ][ HighestTid ]` for the `currLayerId` for the operation point under test when `currLayerId` is greater than 0. Otherwise if operation point DPB information parameters `op_dpb_info_parameters()` are not present for the operation point under test, `MaxDecPicBufferingMinus1[ currLayerId ][ HighestTid ]` is set to `sps_max_dec_pic_buffering_minus1[ HighestTid ]` from the active SPS (when `curr`

rLayerId is equal to 0) or from the active layer SPS for the value of currLayerId.

- [0504] When one or more of the following conditions are true, the "bumping" process specified in subclause F.13.5.2.4 is invoked repeatedly while further decrementing the DPB fullness by one for each additional picture storage buffer that is emptied, until none of the following conditions are true:
- [0505] (1) The number of pictures with nuh\_layer\_id equal to currLayerId in the DPB that are marked as "needed for output" is greater than MaxNumRe-orderPics[CurrLayerId][HighestTid].
- [0506] (2) If MaxLatencyIncreasePlus1[CurrLayerId][HighestTid] is not equal to 0 and there is at least one picture with nuh\_layer\_id equal to currLayerId in the DPB that is marked as "needed for output" for which the associated variable PicLatencyCount[ currLayerId ] is greater than or equal to MaxLatencyPictures[CurrLayerId][HighestTid].
- [0507] (3) The number of pictures with nuh\_layer\_id equal to currLayerId in the DPB is greater than or equal to MaxDecPicBuffering[CurrLayerId][HighestTid].
- [0508] The processes specified in this subclause happen instantaneously when the last decoding unit of access unit n containing the current picture is removed from the CPB.
- [0509] The variable currLayerId is set equal to nuh\_layer\_id of the current decoded picture.
- [0510] For each picture in the DPB that is marked as "needed for output" and that has a nuh\_layer\_id value equal to currLayerId, the associated variable PicLatencyCount[ currLayerId ] is set equal to PicLatencyCount[ currLayerId ] + 1.
- [0511] The current picture is considered as decoded after the last decoding unit of the picture is decoded. The current decoded picture is stored in an empty picture storage buffer in the DPB, and the following applies:
- [0512] (A) If the current decoded picture has PicOutputFlag equal to 1, it is marked as "needed for output" and its associated variable PicLatencyCount[ currLayerId ] is set equal to 0.
- [0513] (B) Otherwise (the current decoded picture has PicOutputFlag equal to 0), it is marked as "not needed for output".
- [0514] The current decoded picture is marked as "used for short-term reference".
- [0515] When one or more of the following conditions are true, the "bumping" process specified in subclause F.13.5.2.4 is invoked repeatedly until none of the following conditions are true.
- [0516] (A) The number of pictures with nuh\_layer\_id equal to currLayerId in the DPB that are marked as "needed for output" is greater than MaxNumRe-orderPics[CurrLayerId][HighestTid].
- [0517] (B) MaxLatencyIncreasePlus1[CurrLayerId][HighestTid] is not equal to 0 and there is at least one picture with nuh\_layer\_id equal to currLayerId in the DPB that is marked as "needed for output" for which the associated variable PicLatencyCount[ cur-

rLayerId ] is greater than or equal to MaxLatencyPictures[CurrLayerId][HighestTid].

- [0518] The "bumping" process consists of the following ordered steps:
- [0519] (A) The pictures that are first for output are selected as the ones having the smallest value of PicOrderCntVal of all pictures in the DPB marked as "needed for output".
- [0520] (B) These pictures are cropped, using the conformance cropping window specified in the active SPS for the picture with nuh\_layer\_id equal to 0 or in the active layer SPS for a nuh\_layer\_id value equal to that of the picture, the cropped pictures are output in ascending order of nuh\_layer\_id, and the pictures are marked as "not needed for output".
- [0521] (C) Each picture storage buffer that contains a picture marked as "unused for reference" and that included one of the pictures that was cropped and output is emptied.
- [0522] The VPS Extension may have additional modifications, if desired.
- [0523] Referring to Figure 40, an additional modification may include the DPB parameters being sent in the VPS extension for output layer sets instead of for operation points, where the oop\_dpb\_info\_parameters(j) are illustrated in Figure 41.
- [0524] The num\_dpb\_info\_parameters specifies the number of oop\_dpb\_parameters( ) syntax structures present in the VPS extension RBSP. num\_dpb\_info\_parameters decoders shall be in the range of 0 to num\_output\_layer\_sets, inclusive.
- [0525] The output\_point\_layer\_set\_idx[ i ] specifies the index, into the list of target output layer sets to which the i th oop\_dpb\_info\_parameters( ) syntax structure in the VPS extension applies.
- [0526] The value of output\_point\_layer\_set\_idx[ i ] should be in the range of 0 to num\_output\_layer\_sets, inclusive. It is requirement of bitstream conformance that output\_point\_layer\_set\_idx [ i ] shall not be equal to output\_point\_layer\_set\_idx [ j ] for any j not equal to i.
- [0527] Referring to Figure 42, the oop\_dpb\_info\_paremters(c) may be further modified, where the syntax in the VPS extension may be as illustrated in Figure 43.
- [0528] Referring to Figure 44, the oop\_dpb\_info\_paremters(c) may be further modified, where the syntax in the VPS extension may be as illustrated in Figure45 or Figure 46.
- [0529] An exemplary alternative for the syntax in VPS extension is that
- [0530] [Table 18]

for(j = 0; j <= vps_max_layer_id; j++ )
oop_dpb_info_parameters(j)

- [0531] may be changed to

- [0532] [Table 19]

for(j = 0; j <= vps_max_layers_minus1; j++ )
oop_dpb_info_parameters(j)

[0533] The vps\_max\_layer\_id specifies the maximum allowed value of nuh\_layer\_id of all NAL units in the CVS. The vps\_max\_layers\_minus1, specifies the maximum number of layers that may be present in the CVS, wherein a layer may e.g. be a spatial scalable layer, a quality scalable layer, a texture view or a depth view.

[0534] Another exemplary alternative for the syntax in VPS extension is that

[0535] [Table 20]

for( j = 0; j <= vps_max_layer_id; j++ )
oop_dpb_info_parameters(j)

[0536] may be changed to

[0537] [Table 21]

for( j = 0; j < numOutputLayers; j++ )
oop_dpb_info_parameters(j)

[0538] where numOutputLayers for the selected output layer set index opIsIdx is derived as:

[0539] [Table 22]

for(k=0, numOutputLayers=0; k<=vps_max_layer_id; k++)
if(output_layer_flag[opIsIdx][k])
targetOpLayerIdList [numOutputLayers++] = layer_id_in_nuh[k]

[0540] Another exemplary alternative for the syntax in VPS extension is that

[0541] [Table 23]

for( j = 0; j <= vps_max_layer_id; j++ )
oop_dpb_info_parameters(j)

[0542] may be changed to

[0543] [Table 24]

for( j = 0; j < numDecodedLayers; j++ )
oop_dpb_info_parameters(j)

[0544] where numOutputLayers for the selected opIsIdx is derived as:

[0545] [Table 25]

for(k=0, numOutputLayers=0; k<=vps_max_layer_id; k++)
if(output_layer_flag[opIsIdx][k])
targetOpLayerIdList [numOutputLayers++] = layer_id_in_nuh[k]

[0546] Then a target decoded layer identifier list targetDLayerIdList and numDecodedLayers for the selected opIsIdx is derived as:

[0547]

[Table 26]

for(m=0, numDecodedLayers=0;m< numOutputLayers;m++) {
for(n=0;n<NumDirectRefLayers[LayerIdInVps[targetOpLayerIdList[m]]];n++) {
rLid=RefLayerId[LayerIdInVps[targetOpLayerIdList[m]]][n]
if(rLid not included in targetDLayerIdList[0,..., numDecodedLayers])
targetDLayerIdList[numDecodedLayers++]=rLid;
}
}

[0548] In one embodiment an additional flag maybe signalled to indicate if oop\_dpb\_information\_parameters are signalled for the particular layer as follows:

[0549] [Table 27]

for(j = 0; j <= vps_max_layer_id; j++) {	
vps_layer_info_present_flag[j]	u(1)
if(vps_layer_info_present_flag)	
oop_dpb_info_parameters(j)	
}	

[0550] The vps\_layer\_info\_present\_flag[j] equal to 1 specifies that oop\_dpb\_info\_parameters are present for the j'th layer for the particular output layer set. vps\_layer\_info\_present\_flag[j] equal to 0 specifies that oop\_dpb\_info\_parameters are not present for the j'th layer for the particular output layer set.

[0551] In another embodiment num\_dpb\_info\_parameters decoders shall be in the range of 0 to 1024, inclusive. In yet another embodiment a different fixed number could be used in place of 1024.

[0552] In an alternative embodiment output\_point\_layer\_set\_idx[ i ] is in the range of 0 to 1023, inclusive.

[0553] Referring to Figure 47, another modified VPS extension and layer\_dpb\_info(i) may be used if the DPB parameters are sent in the VPS extension for each layer independently of output layer sets and operation points.

[0554] Referring to Figure 48, a modified layer\_dpb\_info(i) may be used where the syntax element vps\_max\_sub\_layer\_minus1 signaled from VPS is used for all the layers and is not separately signalled in oop\_dpb\_info\_parameters(id) / op\_dpb\_info\_parameters(



id).

- [0555] Referring to Figure 49A and 49B, an exemplary modified vps\_extension is illustrated. The modified vps extension includes new syntax, namely, max\_sub\_layers\_vps\_predict\_flag[i], max\_sub\_layers\_vps\_minus1[i], num\_dpb\_info\_parameters, output\_point\_layer\_set\_idx[i], oop\_dpb\_maxbuffering\_parameters(i), and layer\_dpb\_info\_parameters(i). num\_output\_layer\_sets specifies the number of layer sets for which output layers are specified with output\_layer\_set\_index[ i ] and output\_layer\_flag[ lsIdx ][ j ]. When not present, the value of num\_output\_layer\_sets is inferred to be equal to 0. A layer set describing output layers is an output layer set.
- [0556] 'max\_sub\_layers\_vps\_predict\_flag'[ i ] equal to 1 specifies that max\_sub\_layers\_vps\_minus1[ i ] is inferred to be equal to max\_sub\_layers\_vps\_minus1[ i - 1 ].
- [0557] max\_sub\_layers\_vps\_predict\_flag[ i ] equal to 0 specifies that max\_sub\_layers\_vps\_minus1[ i ] is explicitly signalled. The value of max\_sub\_layers\_vps\_predict\_flag[ 0 ] is inferred to be equal to 0.
- [0558] 'max\_sub\_layers\_vps\_minus1'[ i ] plus 1 specifies the maximum number of temporal sub-layers that may be present in the CVS for layer with nuh\_layer\_id equal to i. The value of max\_vps\_sub\_layers\_vps\_minus1[ i ] shall be in the range of 0 to 6, inclusive. In some cases max\_sub\_layers\_vps\_minus1[i] is used for the inference of the SPS syntax element sps\_max\_sub\_layers\_minus1. When max\_sub\_layers\_vps\_predict\_flag[ i ] is equal to 1, max\_sub\_layers\_vps\_minus1[ i ] is inferred to be equal to max\_sub\_layers\_vps\_minus1[ i - 1 ]. The value of max\_sub\_layers\_vps\_minus1[ 0 ] is inferred to be equal to vps\_max\_sub\_layers\_minus1.
- [0559] The variable MaxSubLayers[ setId ] for setId in the range of 0 to num\_dpb\_info\_parameters- 1, inclusive, is derived as follows:
- [0560] [Table 28]

for( setId = 0; setId < num_dpb_info_parameters; setId++ ) {
lsIdx =
output_layer_set_idx_minus1[ output_point_layer_set_idx[ setId ] ] + 1
highestLayerId =
layerSetLayerIdList[lsIdx][numLayersInIdList[lsIdx]-1]
MaxSubLayers[setId] =
( max_sub_layers_vps_minus1[highestLayerId]+1 )
}

- [0561] 'num\_dpb\_info\_parameters' specifies the number of oop\_dpb\_maxbuffering\_parameters( i ) syntax structures present in the VPS extension RBSP. num\_dpb\_info\_parameters decoders shall be in the range of 0 to numOutput-LayerSets, inclusive.
- [0562] 'output\_point\_layer\_set\_idx'[ i ] specifies the index, into the list of output layer sets to which the i th oop\_dpb\_maxbuffering\_parameters( i ) syntax structure in the VPS extension applies.
- [0563] The value of output\_point\_layer\_set\_idx[ i ] should be in the range of 0 to numOutputLayerSets, inclusive. It is requirement of bitstream conformance that
- [0564] output\_point\_layer\_set\_idx [ i ] shall not be equal to output\_point\_layer\_set\_idx [ j ] for any j not equal to i.
- [0565] Referring to Figure 50, the oop\_dpb\_maxbuffering\_parameters specifies 'sub\_layer\_vps\_buf\_info\_present\_flag'[ i ], 'max\_vps\_dec\_pic\_buffering\_minus1'[ i ][ j ].
- [0566] 'sub\_layer\_vps\_buf\_info\_present\_flag'[ i ] equal to 1 specifies that max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j ] are present for MaxSubLayers[ i ] sub-layers. sub\_layer\_vps\_buf\_info\_present\_flag[ i ] equal to 0 specifies that the values of max\_vps\_dec\_pic\_buffering\_minus1[ i ][ MaxSubLayers[ i ] - 1 ] apply to all sub-layers.
- [0567] 'max\_vps\_dec\_pic\_buffering\_minus1'[ i ][ j ] plus 1 specifies the maximum required size of the decoded picture buffer for the CVS for layer with nuh\_layer\_id equal to highestLayerId in the output layer set associated with index i in units of picture storage buffers when HighestTid is equal to j. The value of max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j ] shall be in the range of 0 to MaxDpbSize - 1 (as specified in subclause A.4), inclusive. When j is greater than 0, max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j ] shall be greater than or equal to max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j - 1 ] In some cases max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j ] is used for inference of the values of the SPS syntax elements sps\_max\_dec\_pic\_buffering\_minus1[ j ]. When max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j ] is not present for i in the range of 0 to MaxSubLayers[ i ] - 2, inclusive, due to sub\_layer\_vps\_buf\_info\_present\_flag[ i ] being equal to 0, it is inferred to be equal to max\_vps\_dec\_pic\_buffering\_minus1[ i ][ MaxSubLayers[ i ] - 1 ].
- [0568] The value of max\_vps\_dec\_pic\_buffering\_minus1[ 0 ][ j ] for each value of j is inferred to be equal to vps\_max\_dec\_pic\_buffering\_minus1[ j ].
- [0569] Referring to Figure 51, the layer\_dpb\_info\_parameters specifies 'sub\_layer\_vps\_ordering\_info\_present\_flag'[ i ], max\_vps\_num\_reorder\_pics[ i ][ j ], 'max\_vps\_latency\_increase\_plus1'[ i ][ j ].

- [0570] 'sub\_layer\_vps\_ordering\_info\_present\_flag'[ i ] equal to 1 specifies that max\_vps\_num\_reorder\_pics[ i ][ j ] and max\_vps\_latency\_increase\_plus1[ i ][ j ] are present for max\_sub\_layers\_vps\_minus1 + 1 sub-layers.
- [0571] sub\_layer\_vps\_ordering\_info\_present\_flag[ i ] equal to 0 specifies that the values of max\_vps\_num\_reorder\_pics[ i ][ vps\_max\_sub\_layers\_minus1 ] and max\_vps\_latency\_increase\_plus1[ i ][ max\_sub\_layers\_vps\_minus1 ] apply to all sub-layers.
- [0572] 'max\_vps\_num\_reorder\_pics'[ i ][ j ] indicates the maximum allowed number of pictures that can precede any picture in the CVS for layer with nuh\_layer\_id equal to i in decoding order and follow that picture in output order when HighestTid is equal to j. The value of max\_vps\_num\_reorder\_pics[ i ][ j ] shall be in the range of 0 to max\_vps\_dec\_pic\_buffering\_minus1[ i ][ j ], inclusive. When j is greater than 0, max\_vps\_num\_reorder\_pics[ i ][ j ] shall be greater than or equal to max\_vps\_num\_reorder\_pics[ i ][ j-1 ]. In some cases max\_vps\_num\_reorder\_pics[ i ][ j ] is used for inference of the values of the SPS syntax element sps\_max\_num\_reorder\_pics[ j ]. When max\_vps\_num\_reorder\_pics[ i ][ j ] is not present for i in the range of 0 to max\_sub\_layers\_vps\_minus1[ i ] - 1, inclusive, due to sub\_layer\_vps\_ordering\_info\_present\_flag[ i ] being equal to 0, it is inferred to be equal to max\_vps\_num\_reorder\_pics[ i ][ max\_sub\_layers\_vps\_minus1[ i ] ].
- [0573] max\_vps\_latency\_increase\_plus1[ i ][ j ] not equal to 0 is used to compute the value of VpsMaxLatencyPictures[ i ][ j ], which specifies the maximum number of pictures that can precede any picture in the CVS for layer with nuh\_layer\_id equal to i in output order and follow that picture in decoding order when HighestTid is equal to j.
- [0574] When max\_vps\_latency\_increase\_plus1[ i ][ j ] is not equal to 0, the value of VpsMaxLatencyPictures[ i ][ j ] is specified as follows:
- [0575] 
$$\text{VpsMaxLatencyPictures}[ i ][ j ] = \text{max\_vps\_num\_reorder\_pics}[ i ][ j ] + \text{max\_vps\_latency\_increase\_plus1}[ i ][ j ] - 1$$
- [0576] When max\_vps\_latency\_increase\_plus1[ i ][ j ] is equal to 0, no corresponding limit is expressed.
- [0577] The value of max\_vps\_latency\_increase\_plus1[ j ][ k ] shall be in the range of 0 to 2<sup>32</sup> - 2, inclusive. In some cases max\_vps\_latency\_increase\_plus1[ i ][ j ] is used for inference of the values of the SPS syntax elements sps\_max\_latency\_increase\_plus1[ j ]. When max\_vps\_latency\_increase\_plus1[ i ][ j ] is not present for i in the range of 0 to max\_sub\_layers\_vps\_minus1[ i ] - 1, inclusive, due to sub\_layer\_vps\_ordering\_info\_present\_flag[ i ] being equal to 0, it is inferred to be equal to max\_vps\_latency\_increase\_plus1[ i ][ max\_sub\_layers\_vps\_minus1[ i ] ].
- [0578] Referring to Figure 52, an exemplary modified vps\_extension is illustrated. The modified vps extension includes further modification of the syntax in Figure 49 with

- new syntax, namely, `sub_layer_vps_buf_info_predict_flag[ i ]`, `sub_layer_vps_ordering_info_predict_flag[ i ]` which are conditionally signalled. `sub_layer_vps_buf_info_predict_flag[ i ]` equal to 1 specifies that `max_vps_dec_pic_buffering_minus1[ i ][ j ]` is inferred to be equal to `max_vps_dec_pic_buffering_minus1[ i - 1 ][ j ]` for each value of `j`.
- [0579] `sub_layer_vps_buf_info_predict_flag[ i ]` equal to 0 specifies that `max_vps_dec_pic_buffering_minus1[ i ][ j ]` for at least one value of `j` is explicitly signalled. `sub_layer_vps_buf_info_predict_flag[ 0 ]` is inferred to be equal to 0. When not present, `sub_layer_vps_buf_info_predict_flag[ i ]` is inferred to be equal to 0.
- [0580] `sub_layer_vps_ordering_info_predict_flag[ i ]` equal to 1 specifies that the syntax elements `sub_layer_vps_ordering_info_present_flag[ i ]`, `max_vps_num_reorder_pics[ i ][ j ]`, and `max_vps_latency_increase_plus1[ i ][ j ]` are inferred to be equal to `sub_layer_vps_ordering_info_present_flag[ i - 1 ]`, `max_vps_num_reorder_pics[ i - 1 ][ j ]`, and `max_vps_latency_increase_plus1[ i - 1 ][ j ]`, respectively. `sub_layer_vps_ordering_info_predict_flag[ i ]` equal to 0 indicates that the syntax elements `sub_layer_vps_ordering_info_present_flag[ i ]`, `max_vps_num_reorder_pics[ i ][ j ]`, and `max_vps_latency_increase_plus1[ i ][ j ]` are explicitly signalled. When not present, the value of `sub_layer_vps_ordering_info_predict_flag[ i ]` is set equal to 0.
- [0581] Other syntax elements and their semantic meanings for Figure 52 are same as those for figure 49.
- [0582] Referring to Figure 53, an exemplary modified vps\_extension is illustrated. The modified vps extension includes further modification of the syntax in Figure 49. In Figure 53 `oop_dpb_maxbuffering_parameters(i,j)` are signalled for each layer `j` for the particular output layer set `i` compared to Figure 49 which signals a single set of `oop_dpb_maxbuffering_parameters(i)` parameters for the highest layer id in the output layer set `i`.
- [0583] Referring to Figure 54, the `oop_dpb_maxbuffering_parameters` specifies `sub_layer_vps_buf_info_present_flag[ i ][ k ]`, `max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]`.
- [0584] `sub_layer_vps_buf_info_present_flag[ i ][ k ]` equal to 1 specifies that `max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]` are present for `max_sub_layers_vps_minus1[ k ]` sub-layers. `sub_layer_vps_buf_info_present_flag[ i ]` equal to 0 specifies that the values of `max_vps_dec_pic_buffering_minus1[ i ][ k ][ max_sub_layers_vps_minus1[ k ] ]` apply to all sub-layers.
- [0585] `max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]` plus 1 specifies the maximum required size of the decoded picture buffer for the CVS for layer with `nuh_layer_id` equal to `k` for the output layer set associated with index `i` in units of picture storage buffers when `HighestTid` is equal to `j`. The value of

$\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][k][j]$  shall be in the range of 0 to  $\text{MaxDpbSize} - 1$  (as specified in subclause A.4), inclusive.

- [0586] When  $j$  is greater than 0,  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][k][j]$  shall be greater than or equal to  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][k][j-1]$  are used for inference of the values of the SPS syntax elements  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[j]$ .
- [0587] When  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][j]$  is not present for  $i$  in the range of 0 to  $\text{max\_sub\_layers\_vps\_minus1}[k] - 1$ , inclusive, due to  $\text{sub\_layer\_vps\_buf\_info\_present\_flag}[i][k]$  being equal to 0, it is inferred to be equal to  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][k][\text{max\_sub\_layers\_vps\_minus1}[k]]$ .
- [0588] The value of  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][0][j]$  for each value of  $i$  and  $j$  is inferred to be equal to  $\text{vps\_max\_dec\_pic\_buffering\_minus1}[j]$ .
- [0589] Referring to Figure 55 the  $\text{oop\_dpb\_maxbuffering\_parameters}$  specifies  $\text{sub\_layer\_vps\_buf\_info\_present\_flag}[i][k]$ ,  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][k][j]$ . Figure 55 is a variant syntax for  $\text{oop\_dpb\_maxbuffering\_parameters}$  compared to the syntax in Figure 54 for  $\text{oop\_dpb\_maxbuffering\_parameters}$ .
- [0590] The variable  $\text{MaxSubLayers}[\text{setId}][k]$  for  $\text{setId}$  in the range of 0 to  $\text{num\_dpb\_info\_parameters} - 1$ , inclusive, is derived as follows:
- [0591] [Table 29]

for( $\text{setId} = 0; \text{setId} < \text{num\_dpb\_info\_parameters}; \text{setId}++$ ) {
for( $k = 1; k \leq \text{vps\_max\_layers\_minus1}; i++$ ) {
$\text{MaxSubLayers}[\text{setId}][k] = (\text{max\_sub\_layers\_vps\_minus1}[k] + 1)$
}
}

- [0592] In this case the  $\text{oop\_dpb\_maxbuffering\_parameters}$  parameters ( $i, k$ ) will be defined as in the Figure 55.
- [0593] Referring to Figure 56, an exemplary modified  $\text{vps\_extension}$  is illustrated. The modified  $\text{vps}$  extension includes further modification of the syntax in Figure 52 with new syntax. In Figure 53 in this variant  $\text{oop\_dpb\_maxbuffering\_parameters}(i, j)$  are signalled for each layer  $j$  for the particular output layer set  $i$  compared to Figure 52 which signals a single set of  $\text{oop\_dpb\_maxbuffering\_parameters}(i)$  parameters for the highest layer id in the output layer set  $i$ .
- [0594] The  $\text{oop\_dpb\_maxbuffering\_parameters}(I, k)$  are as shown in Figure 57.
- [0595] Referring to Figure 58 the  $\text{oop\_dpb\_maxbuffering\_parameters}$  specifies  $\text{sub\_layer\_vps\_buf\_info\_present\_flag}[i][k]$ ,  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[i][k][j]$ . Figure 58 is a variant syntax for  $\text{oop\_dpb\_maxbuffering\_parameters}$

compared to the syntax in Figure 57 for oop\_dpb\_maxbuffering\_parameters.

[0596] The variable MaxSubLayers[ setId ][ k ] for setId in the range of 0 to num\_dpb\_info\_parameters - 1, inclusive, is derived as follows:

[0597] [Table 30]

for( setId = 0; setId < num_dpb_info_parameters; setId++ ) {
for( k = 1; k <= vps_max_layers_minus1; i++ ) {
MaxSubLayers[setId][k] = ( max_sub_layers_vps_minus1[k ]+1 )
}
}

[0598] In this case the oop\_dpb\_maxbuffering\_parameters parameters ( i, k ) will be defined as in the Figure 58.

[0599] Referring to Figure 59, an exemplary modified vps\_extension is illustrated. The modified vps extension includes further modification of the syntax in Figure 56 with new syntax, namely, an additional flag layer\_dpb\_info\_parameters\_presence\_flag which are conditionally signalled. layer\_dpb\_info\_parameters\_presence\_flag makes signalling of layer\_dpb\_info\_parameters(i) in VPS extension optional.

[0600] layer\_dpb\_info\_parameters\_presence\_flag[ i ] equal to 1 specifies that the syntax elements sub\_layer\_vps\_ordering\_info\_predict\_flag[ i ] and layer\_dpb\_info\_parameters(i) are present for vps\_max\_num\_layers\_minus1 layers. layer\_dpb\_info\_parameters\_presence\_flag[ i ] equal to 0 specifies that the syntax elements sub\_layer\_vps\_ordering\_info\_predict\_flag[ i ] and layer\_dpb\_info\_parameters(i) are not present for vps\_max\_num\_layers\_minus1 layers.

[0601] The oop\_dpb\_maxbuffering\_parameters(i,k) are as shown in Figure 60.

[0602] Referring to Figure 61 the oop\_dpb\_maxbuffering\_parameters specifies sub\_layer\_vps\_buf\_info\_present\_flag[ i ][ k ], max\_vps\_dec\_pic\_buffering\_minus1[ i ][ k ][ j ]. Figure 61 is a variant syntax for oop\_dpb\_maxbuffering\_parameters compared to the syntax in Figure 60 for oop\_dpb\_maxbuffering\_parameters.

[0603] The variable MaxSubLayers[ setId ][ k ] for setId in the range of 0 to num\_dpb\_info\_parameters - 1, inclusive, is derived as follows:

[0604] [Table 31]

for( setId = 0; setId < num_dpb_info_parameters; setId++ ) {
for( k = 1; k <= vps_max_layers_minus1; i++ ) {
MaxSubLayers[setId][k] =( max_sub_layers_vps_minus1[k ]+1 )
}
}

[0605] In this case the oop\_dpb\_maxbuffering\_parameters parameters ( i, k ) will be defined as in the Figure 61.

[0606] An exemplary alternative for the syntax in VPS extension is that

[0607] [Table 32]

for( k = 1; j <= vps_max_layers_minus1; k++ )
oop_dpb_maxbuffering_parameters(k)

[0608] may be changed to

[0609] [Table 33]

for( k = 0; k <= vps_max_layers_minus1; k++ )
oop_dpb_maxbuffering_parameters(k)

[0610] Thus index k could start at 0 instead of at 1.

[0611] An exemplary alternative for the syntax in VPS extension is that

[0612] [Table 34]

for( k = 0; k <= vps_max_layers_minus1; k++ )
oop_dpb_maxbuffering_parameters(k)

[0613] may be changed to

[0614] [Table 35]

for( k = 0; k <= vps_max_layer_id; k++ )
oop_dpb_maxbuffering_parameters(k)

[0615] The vps\_max\_layer\_id specifies the maximum allowed value of nuh\_layer\_id of all NAL units in the CVS. The vps\_max\_layers\_minus1, specifies the maximum number of layers that may be present in the CVS, wherein a layer may e.g. be a spatial scalable layer, a quality scalable layer, a texture view or a depth view.

[0616] Another exemplary alternative for the syntax in VPS extension is that

[0617] [Table 36]

for( k = 1; k <= vps_max_layers_minus1; k++ )
oop_dpb_maxbuffering_parameters(l,k)

[0618] may be changed to

[0619] [Table 37]

for( k = 1; k < numOutputLayers; k++ )
oop_dpb_maxbuffering_parameters(l,k)

[0620] where numOutputLayers is derived as

[0621]

[Table 38]

for( setId = 0; setId < num_dpb_info_parameters; setId++ ) {
lsIdx =
output_layer_set_idx_minus1[ output_point_layer_set_idx[ setId ] ] + 1
numOutputLayers = [ numLayersInIdList[ lsIdx ] - 1 ]
}

[0622] or it could be changed to

[0623] [Table 39]

for( k = 1; k <=
numLayersInIdList[output_layer_set_idx_minus1[output_
point_layer_set_idx[ i ] ]-1; i++ ) {
oop_dpb_maxbuffering_parameters( i, k )

[0624] In one embodiment an additional flag maybe signalled to indicate if  
oop\_dpb\_information\_parameters are signalled for the particular layer as follows:

[0625] [Table 40]

for( k = 1; k <= vps_max_layers_minus1; i++ ) {	
vps_layer_info_present_flag[ k ]	u(1)
oop_dpb_maxbuffering_parameters( i, k )	
}	

[0626] vps\_layer\_info\_present\_flag[ k ] equal to 1 specifies that  
oop\_dpb\_maxbuffering\_parameters are present for the k'th layer for the particular  
output layer set. vps\_layer\_info\_present\_flag[ k ] equal to 0 specifies that  
oop\_dpb\_maxbuffering\_parameters are not present for the k'th layer for the particular  
output layer set.

[0627] The syntax elements of non-VCL NAL units (or their default values for some of the  
syntax elements), required for the HRD, are specified in the semantic subclauses of  
clause 7, Annexes D and E.

[0628] Two types of HRD parameter sets (NAL HRD parameters and VCL HRD pa-  
rameters) are used. The HRD parameter sets are signalled through the hrd\_parameters(  
) syntax structure, which may be part of the SPS syntax structure or the VPS syntax  
structure.

[0629] Multiple tests may be needed for checking the conformance of a bitstream, which is  
referred to as the bitstream under test. For each test, the following steps apply in the  
order listed:

[0630] (1) An output layer set under test, denoted as TargetOpLs is selected. The operation



point referred in TargetOpLs by output\_layer\_set\_idx[] identifies the operation point under test. The output layer identifier list OpLayerIdList of TargetOpLs consists of the list of nuh\_layer\_id values, in increasing order of nuh\_layer\_id values, present in the bitstream subset associated with TargetOp and TargetOpLs, which is a subset of the nuh\_layer\_id values present in the bitstream under test. The OpTid of TargetOp is equal to the highest TemporalId present in the bitstream subset associated with TargetOp.

- [0631] (2) TargetDecLayerIdList is set equal to target decoded layer identifier list targetDecLayerIdList for the selected output layer set TargetOpLs, HighestTid is set equal to OpTid of TargetOp, and the sub-bitstream extraction process as specified in clause 10 is invoked with the bitstream under test, HighestTid, and TargetDecLayerIdList as inputs, and the output is assigned to BitstreamToDecode.
- [0632] (3) The hrd\_parameters( ) syntax structure and the sub\_layer\_hrd\_parameters( ) syntax structure applicable to TargetOp are selected. If TargetDecLayerIdList contains all nuh\_layer\_id values present in the bitstream under test, the hrd\_parameters( ) syntax structure in the active SPS (or provided through an external means not specified in this Specification) is selected. Otherwise, the hrd\_parameters( ) syntax structure in the active VPS (or provided through some external means not specified in this Specification) that applies to TargetOp is selected. Within the selected hrd\_parameters( ) syntax structure, if BitstreamToDecode is a Type I bitstream, the sub\_layer\_hrd\_parameters( HighestTid ) syntax structure that immediately follows the condition "if( vcl\_hrd\_parameters\_present\_flag )" is selected and the variable NalHrdModeFlag is set equal to 0; otherwise (BitstreamToDecode is a Type II bitstream), the sub\_layer\_hrd\_parameters( HighestTid ) syntax structure that immediately follows either the condition "if( vcl\_hrd\_parameters\_present\_flag )" (in this case the variable NalHrdModeFlag is set equal to 0) or the condition "if( nal\_hrd\_parameters\_present\_flag )" (in this case the variable NalHrdModeFlag is set equal to 1) is selected. When BitstreamToDecode is a Type II bitstream and NalHrdModeFlag is equal to 0, all non-VCL NAL units except filler data NAL units, and all leading\_zero\_8bits, zero\_byte, start\_code\_prefix\_one\_3bytes, and trailing\_zero\_8bits syntax elements that form a byte stream from the NAL unit stream (as specified in Annex B), when present, are discarded from BitstreamToDecode, and the remaining bitstream is assigned to BitstreamToDecode.
- [0633] A conforming decoder may fulfil all requirements specified in this subclause.
- [0634] (1) A decoder claiming conformance to a specific profile, tier and level shall be able to successfully decode all bitstreams that conform to the bitstream conformance requirements specified in subclause C.4, in the manner specified in Annex A, provided that all VPSs, SPSs and PPSs referred to in the VCL NAL units, and appropriate

buffering period and picture timing SEI messages are conveyed to the decoder, in a timely manner, either in the bitstream (by non-VCL NAL units), or by external means not specified in this Specification.

- [0635] (2) When a bitstream contains syntax elements that have values that are specified as reserved and it is specified that decoders shall ignore values of the syntax elements or NAL units containing the syntax elements having the reserved values, and the bitstream is otherwise conforming to this Specification, a conforming decoder shall decode the bitstream in the same manner as it would decode a conforming bitstream and shall ignore the syntax elements or the NAL units containing the syntax elements having the reserved values as specified.
- [0636] There are two types of conformance of a decoder: output timing conformance and output order conformance.
- [0637] To check conformance of a decoder, test bitstreams conforming to the claimed profile, tier and level, as specified in subclause C.4 are delivered by a hypothetical stream scheduler (HSS) both to the HRD and to the decoder under test (DUT). All cropped decoded pictures output by the HRD shall also be output by the DUT, each cropped decoded picture output by the DUT shall be a picture with PicOutputFlag equal to 1, and, for each such cropped decoded picture output by the DUT, the values of all samples that are output shall be equal to the values of the samples produced by the specified decoding process.
- [0638] For output timing decoder conformance, the HSS operates as described above, with delivery schedules selected only from the subset of values of SchedSelIdx for which the bit rate and CPB size are restricted as specified in Annex A for the specified profile, tier and level, or with "interpolated" delivery schedules as specified below for which the bit rate and CPB size are restricted as specified in Annex A. The same delivery schedule is used for both the HRD and the DUT.
- [0639] When the HRD parameters and the buffering period SEI messages are present with  $cpb\_cnt\_minus1[ \text{HighestTid} ]$  greater than 0, the decoder shall be capable of decoding the bitstream as delivered from the HSS operating using an "interpolated" delivery schedule specified as having peak bit rate  $r$ , CPB size  $c( r )$ , and initial CPB removal delay
- [0640] [Math.16]
- $$( f(r) \div r )$$
- as follows:

[Table 41]

$$\alpha = (r - \text{BitRate}[\text{SchedSelIdx} - 1]) / (\text{BitRate}[\text{SchedSelIdx}] - \text{BitRate}[\text{SchedSelIdx} - 1]), \quad (\text{C-22})$$

$$c(r) = \alpha * \text{CpbSize}[\text{SchedSelIdx}] + (1 - \alpha) * \text{CpbSize}[\text{SchedSelIdx} - 1], \quad (\text{C-23})$$

$$f(r) = \alpha * \text{InitCpbRemovalDelay}[\text{SchedSelIdx}] * \text{BitRate}[\text{SchedSelIdx}] + (1 - \alpha) * \text{InitCpbRemovalDelay}[\text{SchedSelIdx} - 1] * \text{BitRate}[\text{SchedSelIdx} - 1] \quad (\text{C-24})$$

- [0641] for any SchedSelIdx > 0 and r such that BitRate[ SchedSelIdx - 1 ] <= r <= BitRate[ SchedSelIdx ] such that r and c( r ) are within the limits as specified in Annex A for the maximum bit rate and buffer size for the specified profile, tier and level. The InitCpbRemovalDelay[ SchedSelIdx ] can be different from one buffering period to another and have to be re-calculated.
- [0642] For output timing decoder conformance, an HRD as described above is used and the timing (relative to the delivery time of the first bit) of picture output is the same for both the HRD and the DUT up to a fixed delay.
- [0643] For output order decoder conformance, the following applies:
- [0644] (1) The HSS delivers the bitstream BitstreamToDecode to the DUT "by demand" from the DUT, meaning that the HSS delivers bits (in decoding order) only when the DUT requires more bits to proceed with its processing. This means that for this test, the coded picture buffer of the DUT could be as small as the size of the largest decoding unit.
- [0645] (2) A modified HRD as described below is used, and the HSS delivers the bitstream to the HRD by one of the schedules specified in the bitstream BitstreamToDecode such that the bit rate and CPB size are restricted as specified in Annex A. The order of pictures output shall be the same for both the HRD and the DUT.
- [0646] (3) The HRD CPB size is given by CpbSize[ SchedSelIdx ] as specified in subclause E.2.3, where SchedSelIdx and the HRD parameters are selected as specified in subclause C.1. The DPB size is given by sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] + 1 from the active SPS (when nuh\_layer\_id for the current decoded picture is equal to 0) or from the active layer SPS for the value of nuh\_layer\_id of the current decoded picture.
- [0647] In some cases if output layer sets DPB information parameters oop\_dpb\_maxbuffering\_parameters() are present for the selected output layer set, The DPB size is given by max\_vps\_dec\_pic\_buffering\_minus1[CurrLayerId][HighestTid] where currLayerId is the nuh\_layer\_id of the current decoded picture. Otherwise if output layer sets DPB information parameters oop\_dpb\_maxbuffering\_parameters () are not present for the selected output layer set, the DPB Size is given by sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] + 1 from the active SPS (when nuh\_layer\_id for the current decoded picture is equal to 0) or from the active layer SPS

for the value of nuh\_layer\_id of the current decoded picture.)

- [0648] The removal time from the CPB for the HRD is the final bit arrival time and decoding is immediate. The operation of the DPB of this HRD is as described in subclauses C.5.2 through C.5.2.3.
- [0649] The decoded picture buffer contains picture storage buffers. The number of picture storage buffers for nuh\_layer\_id equal to 0 is derived from the active SPS. The number of picture storage buffers for each non-zero nuh\_layer\_id value is derived from the active layer SPS for that non-zero nuh\_layer\_id value. Each of the picture storage buffers contains a decoded picture that is marked as "used for reference" or is held for future output. The process for output and removal of pictures from the DPB as specified in subclause F.13.5.2.2 is invoked, followed by the invocation of the process for picture decoding, marking, additional bumping, and storage as specified in subclause F.13.5.2.3. The "bumping" process is specified in subclause F.13.5.2.4 and is invoked as specified in subclauses F.13.5.2.2 and F.13.5.2.3.
- [0650] The output and removal of pictures from the DPB before the decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously when the first decoding unit of the access unit containing the current picture is removed from the CPB and proceeds as follows.
- [0651] The decoding process for RPS as specified in subclause 8.3.2 is invoked.
- [0652] (1) If the current picture is an IRAP picture with NoRaslOutputFlag equal to 1 and with nuh\_layer\_id equal to 0 that is not picture 0, the following ordered steps are applied:
  - [0653] (A) The variable NoOutputOfPriorPicsFlag is derived for the decoder under test as follows:
    - [0654] (i) If the current picture is a CRA picture, NoOutputOfPriorPicsFlag is set equal to 1 (regardless of the value of no\_output\_of\_prior\_pics\_flag).
    - [0655] (ii) Otherwise, if the value of pic\_width\_in\_luma\_samples, pic\_height\_in\_luma\_samples, or sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] derived from the active SPS is different from the value of pic\_width\_in\_luma\_samples, pic\_height\_in\_luma\_samples, or sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ], respectively, derived from the SPS active for the preceding picture, NoOutputOfPriorPicsFlag may (but should not) be set to 1 by the decoder under test, regardless of the value of no\_output\_of\_prior\_pics\_flag. Although setting NoOutputOfPriorPicsFlag equal to no\_output\_of\_prior\_pics\_flag is preferred under these conditions, the decoder under test is allowed to set NoOutputOfPriorPicsFlag to 1 in this case.
    - [0656] (iii) Otherwise, NoOutputOfPriorPicsFlag is set equal to no\_output\_of\_prior\_pics\_flag.
  - [0657] (B) The value of NoOutputOfPriorPicsFlag derived for the decoder under test is

applied for the HRD as follows:

- [0658] (i) If NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers in the DPB are emptied without output of the pictures they contain, and the DPB fullness is set equal to 0.
- [0659] (ii) Otherwise (NoOutputOfPriorPicsFlag is equal to 0), all picture storage buffers containing a picture that is marked as "not needed for output" and "unused for reference" are emptied (without output), and all non-empty picture storage buffers in the DPB are emptied by repeatedly invoking the "bumping" process specified in subclause F.13.5.2.4, and the DPB fullness is set equal to 0.
- [0660] (iii) Otherwise (the current picture is not an IRAP picture with NoRaslOutputFlag equal to 1 and with nuh\_layer\_id equal to 0), all picture storage buffers containing a picture which are marked as "not needed for output" and "unused for reference" are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness is decremented by one. The variable currLayerId is set equal to nuh\_layer\_id of the current decoded picture.
- [0661] The variables MaxNumReorderPics[ currLayerId ][ HighestTid ], MaxLatencyIncreasePlus1[ currLayerId ][ HighestTid ], MaxLatencyPictures[ currLayerId ][ HighestTid ], MaxDecPicBufferingMinus1[ currLayerId ][ HighestTid ] are derived as follows:
- [0662] If layer DPB information parameters layer\_dpb\_info\_parameters() are present in VPS, MaxNumReorderPics[ currLayerId ][ HighestTid ] is set to vps\_max\_num\_reorder\_pics[ HighestTid ] when currLayerId is equal to 0 or is set to max\_vps\_num\_reorder\_pics[ CurrLayerId ][ HighestTid ] for the currLayerId when currLayerId is greater than 0. Otherwise if layer DPB information parameters layer\_dpb\_info\_parameters() are not present MaxNumReorderPics[ currLayerId ][ HighestTid ] is set to sps\_max\_num\_reorder\_pics[ HighestTid ] from the active SPS (when currLayerId is equal to 0) or from the active layer SPS for the value of currLayerId.
- [0663] If layer DPB information parameters layer\_dpb\_info\_parameters() are present in VPS, MaxLatencyIncreasePlus1[ currLayerId ][ HighestTid ] is set to vps\_max\_latency\_increase\_plus1[ HighestTid ] when currLayerId is equal to 0 or is set to max\_vps\_latency\_increase\_plus1[ CurrLayerId ][ HighestTid ] for the currLayerId when currLayerId is greater than 0. If layer DPB information parameters layer\_dpb\_info\_parameters() are present in VPS, MaxLatencyPictures[ currLayerId ][ HighestTid ] is set to SpsMaxLatencyPictures [ HighestTid ] when currLayerId is equal to 0 or is set to VpsMaxLatencyPictures [ CurrLayerId ][ HighestTid ] for the currLayerId when currLayerId is greater than 0. Otherwise if layer DPB information parameters layer\_dpb\_info\_parameters() are not present for the for the operation point

under test,  $\text{MaxLatencyIncreasePlus1}[\text{currLayerId}][\text{HighestTid}]$  is set to  $\text{sps\_max\_latency\_increase\_plus1}[\text{HighestTid}]$  of the active SPS (when  $\text{currLayerId}$  is equal to 0) or the active layer SPS for the value of  $\text{currLayerId}$  and  $\text{MaxLatencyPictures}[\text{currLayerId}][\text{HighestTid}]$  is set to  $\text{SpsMaxLatencyPictures}[\text{currLayerId}][\text{HighestTid}]$  derived from the active SPS (when  $\text{currLayerId}$  is equal to 0) or from the active layer SPS for the value of  $\text{currLayerId}$ .

- [0664] If operation point DPB information parameters  $\text{oop\_dpb\_maxbuffering\_parameters}()$  are present for the selected output layer set,  $\text{MaxDecPicBufferingMinus1}[\text{currLayerId}][\text{HighestTid}]$  is set to  $\text{vps\_max\_dec\_pic\_buffering\_minus1}[\text{HighestTid}]$  when  $\text{currLayerId}$  is equal to 0 or is set to  $\text{max\_vps\_dec\_pic\_buffering\_minus1}[\text{CurrLayerId}][\text{HighestTid}]$  for the  $\text{currLayerId}$  for the operation point under test when  $\text{currLayerId}$  is greater than 0. Otherwise if operation point DPB information parameters  $\text{oop\_dpb\_maxbuffering\_parameters}()$  are not present for the operation point under test,  $\text{MaxDecPicBufferingMinus1}[\text{currLayerId}][\text{HighestTid}]$  is set to  $\text{sps\_max\_dec\_pic\_buffering\_minus1}[\text{HighestTid}]$  from the active SPS (when  $\text{currLayerId}$  is equal to 0) or from the active layer SPS for the value of  $\text{currLayerId}$ .
- [0665] When one or more of the following conditions are true, the "bumping" process specified in subclause F.13.5.2.4 is invoked repeatedly while further decrementing the DPB fullness by one for each additional picture storage buffer that is emptied, until none of the following conditions are true:
- [0666] (1) The number of pictures with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB that are marked as "needed for output" is greater than  $\text{MaxNumReorderPics}[\text{CurrLayerId}][\text{HighestTid}]$ .
- [0667] (2) If  $\text{MaxLatencyIncreasePlus1}[\text{CurrLayerId}][\text{HighestTid}]$  is not equal to 0 and there is at least one picture with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB that is marked as "needed for output" for which the associated variable  $\text{PicLatencyCount}[\text{currLayerId}]$  is greater than or equal to  $\text{MaxLatencyPictures}[\text{CurrLayerId}][\text{HighestTid}]$ .
- [0668] (3) The number of pictures with  $\text{nuh\_layer\_id}$  equal to  $\text{currLayerId}$  in the DPB is greater than or equal to  $\text{MaxDecPicBuffering}[\text{CurrLayerId}][\text{HighestTid}]$ .
- [0669] The processes specified in this subclause happen instantaneously when the last decoding unit of access unit  $n$  containing the current picture is removed from the CPB.
- [0670] The variable  $\text{currLayerId}$  is set equal to  $\text{nuh\_layer\_id}$  of the current decoded picture.
- [0671] For each picture in the DPB that is marked as "needed for output" and that has a  $\text{nuh\_layer\_id}$  value equal to  $\text{currLayerId}$ , the associated variable  $\text{PicLatencyCount}[\text{currLayerId}]$  is set equal to  $\text{PicLatencyCount}[\text{currLayerId}] + 1$ .
- [0672] The current picture is considered as decoded after the last decoding unit of the picture is decoded. The current decoded picture is stored in an empty picture storage buffer in the DPB, and the following applies:

- [0673] (A) If the current decoded picture has PicOutputFlag equal to 1, it is marked as "needed for output" and its associated variable PicLatencyCount[ currLayerId ] is set equal to 0.
- [0674] (B) Otherwise (the current decoded picture has PicOutputFlag equal to 0), it is marked as "not needed for output".
- [0675] The current decoded picture is marked as "used for short-term reference".
- [0676] When one or more of the following conditions are true, the "bumping" process specified in subclause F.13.5.2.4 is invoked repeatedly until none of the following conditions are true.
- [0677] (A) The number of pictures with nuh\_layer\_id equal to currLayerId in the DPB that are marked as "needed for output" is greater than MaxNumReorderPics[CurrLayerId][HighestTid].
- [0678] (B) MaxLatencyIncreasePlus1[CurrLayerId][HighestTid] is not equal to 0 and there is at least one picture with nuh\_layer\_id equal to currLayerId in the DPB that is marked as "needed for output" for which the associated variable PicLatencyCount[ currLayerId ] is greater than or equal to MaxLatencyPictures[CurrLayerId][HighestTid].
- [0679] The "bumping" process consists of the following ordered steps:
- [0680] (A) The pictures that are first for output are selected as the ones having the smallest value of PicOrderCntVal of all pictures in the DPB marked as "needed for output".
- [0681] (B) These pictures are cropped, using the conformance cropping window specified in the active SPS for the picture with nuh\_layer\_id equal to 0 or in the active layer SPS for a nuh\_layer\_id value equal to that of the picture, the cropped pictures are output in ascending order of nuh\_layer\_id, and the pictures are marked as "not needed for output".
- [0682] (C) Each picture storage buffer that contains a picture marked as "unused for reference" and that included one of the pictures that was cropped and output is emptied.
- [0683] Referring to FIG. 62, an exemplary modified sequence parameter set (sps) syntax seq\_parameter\_set\_rbsp is illustrated. The modified sps includes a sps\_dpb\_params\_present\_flag. In one embodiment based on the value of this flag and based on the value of nuh\_layer\_id some of the syntax elements (e.g. sps\_sub\_layer\_ordering\_info\_present\_flag, sps\_max\_dec\_pic\_buffering\_minus1[i], sps\_max\_num\_reorder\_pics[i], sps\_max\_latency\_increase\_plus1[i] may not be signaled.
- [0684] sps\_dpb\_params\_present\_flag equal to 0 specifies that, for SPS with nuh\_layer\_id > 0 the syntax elements sps\_sub\_layer\_ordering\_info\_present\_flag, sps\_max\_dec\_pic\_buffering\_minus1[ i ], sps\_max\_num\_reorder\_pics[ i ], sps\_max\_latency\_increase\_plus1[ i ] are not present in this SPS. The value of these pa-

parameters is set equal to the value of `sub_layer_vps_buf_info_present_flag[i]`, `max_vps_num_reorder_pics[i][j]` and `max_vps_latency_increase_plus1[i][j]` parameters signaled in active VPS extension. `sps_dpb_params_present_flag` equal to 1 specifies that, for SPS with `nuh_layer_id > 0` the syntax elements `sps_sub_layer_ordering_info_present_flag`, `sps_max_dec_pic_buffering_minus1[i]`, `sps_max_num_reorder_pics[i]`, `sps_max_latency_increase_plus1[i]` are present in this SPS.

- [0685] In another embodiment one or more of the syntax elements may be signaled using a known fixed number of bits instead of `u(v)` instead of `ue(v)`. For example they could be signaled using `u(8)` or `u(16)` or `u(32)` or `u(64)`, etc.
- [0686] In another embodiment one or more of these syntax element could be signaled with `ue(v)` or some other coding scheme instead of fixed number of bits such as `u(v)` coding.
- [0687] In another embodiment the names of various syntax elements and their semantics may be altered by adding a plus1 or plus2 or by subtracting a minus1 or a minus2 compared to the described syntax and semantics.
- [0688] In yet another embodiment various syntax elements may be signaled per picture anywhere in the bitstream. For example they may be signaled in slice segment header, pps/ sps/ vps/ or any other parameter set or other normative part of the bitstream.
- [0689] In yet another embodiments all the concepts defined in this invention related to output layer sets could be applied to output operation points [2,3] and/ or to operation points [1].

### Example 3

- [0690] A method for video coding is disclosed. The method includes beginning to parse a first slice header of a current picture. It is determined which steps performed by a decoded picture buffer (DPB) will be picture based and which steps will be access unit (AU) based. A removal from the DPB is performed. A picture output from the DPB is performed. A decoding and storing of a current decoded picture in the DPB are performed. The current decoded picture in the DPB is marked. An additional picture output from the DPB is also performed.
- [0691] In some configurations, the removal and the output from the DPB may be based on at least one AU output flag, such as an AU output flag, an AU no rasl output flag and/or an AU no output of prior pictures flag. The AU output flag may be derived based on syntax elements signaled in the bitstream and other conditions. The AU flags represent flags derived and applied at access unit (AU) level. In some cases, these will be flags different than similar flags which are signaled or derived at picture level, such as a picture output flag (e.g., `pic_output_flag`), a picture no rasl output flag (e.g., `No-RaslOutputFlag`) and/or a picture no output of prior pictures flag (e.g.,



no\_output\_of\_prior\_pics flag).

- [0692] In one configuration, a first AU output flag may be derived and used by the DPB and a second AU output flag is derived and used by the DPB. In another configuration, all picture output flag syntax element values may be constrained to the same value for all coded pictures in an AU when a picture output flag is present in the first slice header. In some configurations, all picture output flag syntax element values may be constrained to the same value for all coded pictures in an AU when a no picture output flag is present in the first slice header.
- [0693] In one configuration, the removal may be picture based, the picture output may be AU (access unit) based, the storing and decoding may be picture based, the marking may be picture based, and the additional picture output may be AU based. The removal from the DPB may remove one or more pictures from the DPB before decoding of the current picture.
- [0694] In another configuration, the removal may be picture based, and the picture output, the decoding and storing, the marking and the additional picture output may be AU based. In yet another configuration, the removal, the picture output, the decoding and storing, the marking and the additional picture output may be AU based. In another configuration, the removal, the decoding and storing and the marking may be picture based, and the picture output and the additional picture output may be AU based. In yet another configuration, the removal, the picture output, the decoding and storing, the marking and the additional picture output may be picture based.
- [0695] Marking the current decoded picture in the DPB may include a reference marking step and an output marking step. The reference marking step may be picture based and the output marking step may be AU based. A DPB fullness may be incremented by one when a decoded picture is stored in the DPB in an empty storage buffer. The DPB fullness may be decremented by one when a picture is output from the DPB. The DPB fullness may be tracked per layer. The DPB fullness may also be tracked for an output layer set.
- [0696] The DPB may include separately identified and managed picture buffers for decoded pictures having one or more of different resolutions, different bit-depths and different color chromaticity. The DPB may include a common pool of picture storage buffers. A decoded picture may be stored in the picture storage buffers based on at least one of size, resolution and bit-depth. In one configuration, a decoded picture may be stored in one picture buffer slot of the picture storage buffers.
- [0697] The method may be performed by a decoder within an electronic device that conforms to a scalable high efficiency video coding (SHVC) standard. The method may also be performed by a decoder within an electronic device that conforms to a multi-view high efficiency video coding (MV-HEVC) standard.

- [0698] An electronic device configured for video coding is also disclosed. The electronic device includes a processor and memory in electronic communication with the processor. Instructions in the memory are executable to begin to parse a first slice header of a current picture. Instructions in the memory are also executable to determine which steps performed by a decoded picture buffer (DPB) will be picture based and which steps will be access unit (AU) based. Instructions in the memory are further executable to perform a removal from the DPB. Instructions in the memory are also executable to perform a picture output from the DPB. Instructions in the memory are further executable to perform a decoding and storing of a current decoded picture in the DPB. Instructions in the memory are also executable to mark the current decoded picture in the DPB. Instructions in the memory are further executable to perform an additional picture output from the DPB.
- [0699] Various configurations are now described with reference to the Figures, where like reference numbers may indicate functionally similar elements. The systems and methods as generally described and illustrated in the Figures herein could be arranged and designed in a wide variety of different configurations. Thus, the following more detailed description of several configurations, as represented in the Figures, is not intended to limit scope, as claimed, but is merely representative of the systems and methods.
- [0700] Figure 63 is a block diagram illustrating video coding between multiple electronic devices 2102a-b. A first electronic device 2102a and a second electronic device 2102b are illustrated. However, it should be noted that one or more of the features and functionality described in relation to the first electronic device 2102a and the second electronic device 2102b may be combined into a single electronic device 102 in some configurations. Each electronic device 102 may be configured to encode video and/or decode video. The electronic devices 102 may be configured to use hybrid decoded picture buffer (DPB) operation. Hybrid decoded picture buffer (DPB) operation refers to scenarios where the various steps of removal, output (bumping), storage, marking and additional output (bumping) performed on a decoded picture buffer (DPB) 116 happen on either a picture basis or an access unit (AU) basis. Although specific combinations of these steps are referred to as being performed on a picture basis or an access unit (AU) basis, all possible combinations of doing each of these steps individually on either a picture basis or an access unit (AU) basis are supported.
- [0701] As used herein, access unit (AU) refers to a set of network abstraction layer (NAL) units that are associated with each other according to a specified classification rule, that are consecutive in decoding order, and that include the video coding layer (VCL) NAL units of all coded pictures associated with the same output time and their associated non-VCL NAL units. The base layer is a layer in which all VCL NAL units

have a nuh\_layer\_id equal to 0. A coded picture is a coded representation of a picture that includes VCL NAL units with a particular value of nuh\_layer\_id and that includes all the coding tree units of the picture. In some cases, a coded picture may be called a layer component. Additional details about steps being picture based or access unit (AU) based are given in relation to Figures 69 and 70 below.

- [0702] In one configuration, each of the electronic devices 102 may conform to the High Efficiency Video Coding (HEVC) standard, the Scalable High Efficiency Video Coding (SHVC) standard or the Multi-view High Efficiency Video Coding (MV-HEVC) standard. The HEVC standard is a video compression standard that acts as a successor to H.264/MPEG-4 AVC (Advanced Video Coding) and that provides improved video quality and increased data compression ratios. As used herein, a picture is an array of luma samples in monochrome format or an array of luma samples and two corresponding arrays of chroma samples in 4:2:0, 4:2:2 and 4:4:4 colour format or some other colour format. The operation of a hypothetical reference decoder (HRD) and the operation of the output order decoded picture buffer (DPB) 116 are described for SHVC and MV-HEVC in JCTVC-M1008, JCTVC-L1008, JCTVC-D1004, JCT3V-C1004, JCTVC-L0453 and JCTVC-L0452.
- [0703] The first electronic device 2102a may include an encoder 2108 and an overhead signaling module 2112. The first electronic device 2102a may obtain an input picture 2106. In some configurations, the input picture 2106 may be captured on the first electronic device 2102a using an image sensor, retrieved from memory and/or received from another electronic device 102. The encoder 2108 may encode the input picture 2106 to produce encoded data 2110. For example, the encoder 2108 may encode a series of input pictures 2106 (e.g., video). The encoded data 2110 may be digital data (e.g., a bitstream).
- [0704] The overhead signaling module 2112 may generate overhead signaling based on the encoded data 2110. For example, the overhead signaling module 2112 may add overhead data to the encoded data 2110 such as slice header information, video parameter set (VPS) information, sequence parameter set (SPS) information, picture parameter set (PPS) information, picture order count (POC), reference picture designation, etc. In some configurations, the overhead signaling module 2112 may produce a wrap indicator that indicates a transition between two sets of pictures.
- [0705] The encoder 2108 (and overhead signaling module 2112, for example) may produce a bitstream 2114. The bitstream 2114 may include encoded picture data based on the input picture 2106. In some configurations, the bitstream 2114 may also include overhead data, such as slice header information, VPS information, SPS information, PPS information, etc. As additional input pictures 2106 are encoded, the bitstream 2114 may include one or more encoded pictures. For instance, the bitstream 2114 may

include one or more encoded reference pictures and/or other pictures.

- [0706] The bitstream 2114 may be provided to a decoder 2104. In one example, the bitstream 2114 may be transmitted to the second electronic device 2102b using a wired or wireless link. In some cases, this may be done over a network, such as the Internet or a Local Area Network (LAN). As illustrated in Figure 63, the decoder 2104 may be implemented on the second electronic device 2102b separately from the encoder 2108 on the first electronic device 2102a. However, it should be noted that the encoder 2108 and decoder 2104 may be implemented on the same electronic device 102 in some configurations. When the encoder 2108 and decoder 2104 are implemented on the same electronic device, for instance, the bitstream 2114 may be provided over a bus to the decoder 2104 or stored in memory for retrieval by the decoder 2104.
- [0707] The decoder 2104 may receive (e.g., obtain) the bitstream 2114. The decoder 2104 may generate a decoded picture 2118 (e.g., one or more decoded pictures 2118) based on the bitstream 2114. The decoded picture 2118 may be displayed, played back, stored in memory and/or transmitted to another device, etc.
- [0708] The decoder 2104 may include a decoded picture buffer (DPB) 116. The decoded picture buffer (DPB) 116 may be a buffer holding decoded pictures for reference, output reordering or output delay specified for a hypothetical reference decoder (HRD). On an electronic device 102, a decoded picture buffer (DPB) 116 may be used to store reconstructed (e.g., decoded) pictures at a decoder 2104. These stored pictures may then be used, for example, in an inter-prediction mechanism. When pictures are decoded out of order, the pictures may be stored in the decoded picture buffer (DPB) 116 so they can be displayed later in order.
- [0709] In the H.264 or advance video coding (AVC) standard, decoded picture buffer (DPB) 116 management (e.g., deletion, addition of pictures, reordering of pictures, etc.) is carried out using memory management control operations (MMCO). Many different decoded picture buffer (DPB) 116 management approaches are under consideration.
- [0710] The decoder 2104 may include a hybrid decoded picture buffer (DPB) operation module 2120. The hybrid decoded picture buffer (DPB) operation module 2120 may allow for decoded picture buffer (DPB) 116 management approaches with picture basis steps 2122 and/or decoded picture buffer (DPB) 116 management approaches with access unit (AU) based steps 2124. For example, one advantage of the use of picture based steps 2124 for removal, storage and reference marking is that optimal decoded picture buffer (DPB) 116 memory will be used by various layers. Thus, the overall required memory when using picture based steps may be lower. One advantage of the use of access unit (AU) based steps 2124 for output (including output, output marking and additional output) is that the output process may be simplified.
- [0711] Figure 64 is a flow diagram of a method 2200 for hybrid decoded picture buffer

(DPB) 116 operation. The method 2200 may be performed by a decoded picture buffer (DPB) 116 as part of a decoder 2104 on an electronic device 102. In one configuration, the method 2200 may be performed by a hybrid decoded picture buffer (DPB) operation module 2120. The decoded picture buffer (DPB) 116 may begin parsing 2202 a first slice header of a current picture. The decoded picture buffer (DPB) 116 may determine 2204 which steps of the hybrid decoded picture buffer (DPB) operation will be picture based and which steps will be access unit (AU) based.

[0712] The decoded picture buffer (DPB) 116 may perform 2206 a removal (without output) from the decoded picture buffer (DPB) 116. The removal may remove pictures from the decoded picture buffer (DPB) 116 before the decoding of the current picture. The decoded picture buffer (DPB) 116 may perform 2208 a picture output (bumping) from the decoded picture buffer (DPB) 116. A picture output (bumping) may refer to the output of pictures from the decoded picture buffer (DPB) 116 at the coded picture buffer (CPB) removal time. In some configurations, the term bumping may be used to indicate the output of one or more pictures from the decoded picture buffer (DPB) 116. Thus, the terms bumping and output may be used interchangeably.

[0713] The decoded picture buffer (DPB) 116 may decode 2210 and store the current picture in the decoded picture buffer (DPB) 116. The decoded picture buffer (DPB) 116 may mark 2212 the current decoded picture stored in the decoded picture buffer (DPB) 116. For example, the decoded picture buffer (DPB) 116 may mark 2212 the current decoded picture as "unused for reference," "used for reference," "needed for output" or "not needed for output." The decoded picture buffer (DPB) 116 may also perform 2214 another picture output (additional bumping) from the decoded picture buffer (DPB) 116. In some configurations, a repeated removal/bumping of pictures from the decoded picture buffer (DPB) 116 may occur until certain conditions are satisfied.

[0714] Figure 65 is a flow diagram of another method 2300 for hybrid decoded picture buffer (DPB) 116 operation. For example, the method 2300 of Figure 65 may be a preferred method for hybrid decoded picture buffer (DPB) 116 operation. The method 2300 may be performed by a decoded picture buffer (DPB) 116 as part of a decoder 2104 on an electronic device 102. In one configuration, the method 2300 may be performed by a hybrid decoded picture buffer (DPB) operation module 2120. The decoded picture buffer (DPB) 116 may begin parsing 2302 a first slice header of a current picture. The term hybrid refers to the fact that some of the decoded picture buffer (DPB) 116 operation steps are performed on a picture basis and some of the decoded picture buffer (DPB) 116 operation steps are performed on an access unit (AU) basis. The decoded picture buffer (DPB) 116 may perform 2304 a picture based removal (without output) from the decoded picture buffer (DPB) 116. The decoded picture buffer (DPB) 116 may perform 2306 an access unit (AU) based picture output

(bumping) from the decoded picture buffer (DPB) 116. The decoded picture buffer (DPB) 116 may perform 2308 picture based decoding and storage of the current picture in the decoded picture buffer (DPB).

- [0715] The decoded picture buffer (DPB) 116 may perform 310 picture based marking of the current decoded picture in the decoded picture buffer (DPB). The marking step performed by the decoded picture buffer (DPB) 116 may be further sub-divided to include both a reference marking step and an output marking step. As used herein, marking pictures as "unused for reference" or "used for reference" is referred to as the reference marking step. A decoded picture in the decoded picture buffer (DPB) 116 can be marked as only one of "unused for reference," "used for short-term reference," or "used for long-term reference" at any given moment during the decoding process operation. Assigning one of these markings to a picture implicitly removes another of the markings that is assigned to the picture. When a picture is referred to as being marked as "used for reference," this collectively refers to the picture being marked as either "used for short-term reference" or "used for long-term reference," but never both. As used herein, marking pictures as "needed for output" or "not needed for output" is referred to as the output marking step.
- [0716] The decoded picture buffer (DPB) 116 may operate in such a manner that the reference marking step and the output marking step could happen on either a picture basis or on an access unit (AU) basis. In general, all possible combinations (typically four combinations) of doing these two marking steps individually on a picture basis or on an access unit (AU) basis are supported. However, it may be preferable that the reference marking step is picture based and the output marking step is access unit (AU) based.
- [0717] The decoded picture buffer (DPB) 116 may also perform 2312 an access unit (AU) based output marking of the current decoded picture. The decoded picture buffer (DPB) 116 may perform 2314 another access unit (AU) based picture output (additional bumping) from the decoded picture buffer (DPB) 116. In some configurations, a repeated removal/bumping of pictures from the decoded picture buffer (DPB) 116 may occur until certain conditions are satisfied.
- [0718] In some approaches, bitstream conformance constraints may be required for flags in the slice segment header. In some cases, the constraints may be applied for flags across pictures belonging to the same access unit (AU). For example, the flags `pic_output_flag` and/or `no_output_of_prior_pics_flag` may be required to follow bitstream conformance constraints. For instance, JCTVC-L1003, JCTVC-M1008 and JCT3V-D1004 each describe signaling in the slice segment header using `pic_output_flag` and `no_output_of_prior_pics_flag`. Further, flags such as `PicOutputFlag`, `NoRaslOutputFlag` and `NoOutputOfPriorPicsFlag` are derivable based on

the syntax elements and NAL unit types.

- [0719] JCTVC-L1003, JCTVC-M1008, and JCT3V-D1004, also include descriptions of DPB for HEVC, SHVC and MV-HEVC. JCTVC-M1008 SHVC Draft Text 1 provides text draft for a scalable extension of HEVC. JCT3V-D1004 MV-HEVC Draft Text 4 describes text draft for multi-view extension of HEVC.
- [0720] The video coded bitstream, according to JCTVC-L1003, JCTVC-M1008 and/or JCT3V-D1004, may include a syntax structure that is placed into logical data packets generally referred to as Network Abstraction Layer (NAL) units. Each NAL unit includes a NAL unit header, such as a two-byte NAL unit header (e.g., 16 bits), to identify the purpose of the associated data payload. For example, each coded slice (and/or picture) may be coded in one or more slice (and/or picture) NAL units. Other NAL units may be included for other categories of data, such as for example, supplemental enhancement information, a coded slice of temporal sub-layer access (TSA) picture, a coded slice of step-wise temporal sub-layer access (STSA) picture, a coded slice a non-TSA, non-STSA trailing picture, a coded slice of broken link access picture, a coded slice of instantaneous decoded refresh picture, a coded slice of clean random access picture, a coded slice of decodable leading picture, a coded slice of tagged for discard picture, a video parameter set, sequence parameter set, a picture parameter set, access unit delimiter, an end of sequence, an end of bitstream, filler data and/or a sequence enhancement information message. Table (7) below illustrates one example of NAL unit codes and NAL unit type classes. Other NAL unit types may also be included, as desired.
- [0721] It should also be understood that the NAL unit type values for the NAL units shown in the Table (7) may be rearranged and reassigned. Also, additional NAL unit types may be added or removed.
- [0722] An intra random access point (IRAP) picture is a coded picture for which each video coding layer NAL unit has nal\_unit\_type in the range of BLA\_W\_LP to RSV\_IRAP\_VCL23, inclusive, as shown in Table (7) below. An IRAP picture includes only Intra coded (I) slices.
- [0723] An instantaneous decoding refresh (IDR) picture is an IRAP picture for which each video coding layer NAL unit has nal\_unit\_type equal to IDR\_W\_RADL or IDR\_N\_LP, as shown in Table (7). An instantaneous decoding refresh (IDR) picture includes only I slices and may be the first picture in the bitstream in decoding order, or may appear later in the bitstream.
- [0724] Each IDR picture is the first picture of a coded video sequence (CVS) in decoding order. A broken link access (BLA) picture is an IRAP picture for which each video coding layer NAL unit has nal\_unit\_type equal to BLA\_W\_LP, BLA\_W\_RADL, or BLA\_N\_LP, as shown in Table (7).

[0725] A BLA picture includes only I slices, and may be the first picture in the bitstream in decoding order, or may appear later in the bitstream. Each BLA picture begins a new coded video sequence, and has the same effect on the decoding process as an IDR picture. However, a BLA picture includes syntax elements that specify a non-empty reference picture set.

[0726] [Table 42]

nal_unit_type	Name of nal_unit_type	Content of NAL unit and raw byte sequence payload (Rbsp) syntax structure	NAL unit type class
0 1	TRAIL_N TRAIL_R	Coded slice segment of a non-TSA, non-STSA trailing picture slice segment layer rbsp()	Video Coding Layer (VCL)
2 3	TSA_N TSA_R	Coded slice segment of a temporal sub-layer access (TSA) picture slice segment layer rbsp()	VCL
4 5	STSA_N STSA_R	Coded slice segment of an Step-wise Temporal sub-layer access (STSA) picture slice segment layer rbsp()	VCL
6 7	RADL_N RADL_R	Coded slice segment of a random access decodable leading (RADL) picture slice segment layer rbsp()	VCL
8 9	RASL_N RASL_R	Coded slice segment of a random access skipped leading (RASL) picture slice segment layer rbsp()	VCL
10 12 14	RSV_VCL_N10 RSV_VCL_N12 RSV_VCL_N14	Reserved non-IRAP sub-layer non-reference VCL NAL unit types	VCL
11 13 15	RSV_VCL_R11 RSV_VCL_R13 RSV_VCL_R15	Reserved non-IRAP sub-layer reference VCL NAL unit types	VCL
16 17 18	BLA_W_LP BLA_W_RADL BLA_N_LP	Coded slice segment of a broken link access (BLA) picture slice segment layer rbsp()	VCL
19 20	IDR_W_RADL IDR_N_LP	Coded slice segment of an instantaneous decoding refresh (IDR) picture slice segment layer rbsp()	VCL
21	CRA_NUT	Coded slice segment of a clean random access (CRA) picture slice segment layer rbsp()	VCL
22 23	RSV_IRAP_VCL22 RSV_IRAP_VCL23	Reserved IRAP VCL NAL unit types	VCL
24..31	RSV_VCL24.. RSV_VCL31	Reserved non-IRAP VCL NAL unit types	VCL
32	VPS_NUT	Video parameter set video_parameter_set_rbsp()	non-video coding layer (non-VCL)
33	SPS_NUT	Sequence parameter set seq_parameter_set_rbsp()	non-VCL
34	PPS_NUT	Picture parameter set pic_parameter_set_rbsp()	non-VCL
35	AUD_NUT	Access unit delimiter access_unit_delimiter_rbsp()	non-VCL
36	EOS_NUT	End of sequence end_of_seq_rbsp()	non-VCL
37	EOB_NUT	End of bitstream end_of_bitstream_rbsp()	non-VCL
38	FD_NUT	Filler data filler_data_rbsp()	non-VCL
39 40	PREFIX_SEI_NUT SUFFIX_SEI_NUT	Supplemental enhancement information sei_rbsp()	non-VCL
41..47	RSV_NVCL41.. RSV_NVCL47	Reserved	non-VCL
48..63	UNSPEC48.. UNSPEC63	Unspecified	non-VCL

Table (7)

[0727] Referring to Table (8) below, the NAL unit header syntax may include two bytes of data, namely, 16 bits. The first bit may be a "forbidden\_zero\_bit" that is always set to



zero at the start of a NAL unit. The next six bits may be a "nal\_unit\_type" which specifies the type of raw byte sequence payloads ("RBSP") data structure included in the NAL unit as shown in Table (7) above. The next 6 bits may be a "nuh\_layer\_id" which specify the identifier of the layer. In some cases, these six bits may be specified as "nuh\_reserved\_zero\_6bits" instead. The nuh\_reserved\_zero\_6bits may be equal to 0 in the base specification of the standard. In a scalable video coding and/or syntax extensions, nuh\_layer\_id may specify that this particular NAL unit belongs to the layer identified by the value of these 6 bits.

[0728] The next syntax element may be "nuh\_temporal\_id\_plus1". The nuh\_temporal\_id\_plus1 minus 1 may specify a temporal identifier for the NAL unit. The variable temporal identifier TemporalId may be specified as TemporalId = nuh\_temporal\_id\_plus1 - 1. The temporal identifier TemporalId is used to identify a temporal sub-layer. The variable HighestTid identifies the highest temporal sub-layer to be decoded.

[0729] [Table 43]

nal_unit_header() {	Descriptor
forbidden_zero_bit	f(1)
nal_unit_type	u(6)
nuh_layer_id	u(6)
nuh_temporal_id_plus1	u(3)
}	

Table (8)

[0730] Table (9) below shows an exemplary sequence parameter set (SPS) syntax structure. pic\_width\_in\_luma\_samples specifies the width of each decoded picture in units of luma samples. pic\_width\_in\_luma\_samples shall not be equal to 0. pic\_height\_in\_luma\_samples specifies the height of each decoded picture in units of luma samples. pic\_height\_in\_luma\_samples shall not be equal to 0.

[0731] sps\_max\_sub\_layers\_minus1 plus 1 specifies the maximum number of temporal sub-layers that may be present in each CVS referring to the SPS. The value of sps\_max\_sub\_layers\_minus1 ranges from 0 to 6, inclusive.

[0732] sps\_sub\_layer\_ordering\_info\_present\_flag flag equal to 1 specifies that sps\_max\_dec\_pic\_buffering\_minus1[ i ], sps\_max\_num\_reorder\_pics[ i ], and sps\_max\_latency\_increase\_plus1[ i ] syntax elements are present for sps\_max\_sub\_layers\_minus1 + 1 sub-layers. sps\_sub\_layer\_ordering\_info\_present\_flag equal to 0 specifies that the values of sps\_max\_dec\_pic\_buffering\_minus1[ sps\_max\_sub\_layers\_minus1 ], sps\_max\_num\_reorder\_pics[ sps\_max\_sub\_layers\_minus1 ], and sps\_max\_latency\_increase\_plus1[ sps\_max\_sub\_layers\_minus1 ] apply to all sub-

layers.

- [0733] `sps_max_dec_pic_buffering_minus1[ i ]` plus 1 specifies the maximum required size of the decoded picture buffer for the CVS in units of picture storage buffers when `HighestTid` is equal to `i`. The value of `sps_max_dec_pic_buffering_minus1[ i ]` ranges from 0 to `MaxDpbSize - 1`, inclusive where `MaxDpbSize` specifies the maximum decoded picture buffer size in units of picture storage buffers. When `i` is greater than 0, `sps_max_dec_pic_buffering_minus1[ i ]` shall be greater than or equal to `sps_max_dec_pic_buffering_minus1[ i - 1 ]`. When `sps_max_dec_pic_buffering_minus1[ i ]` is not present for `i` in the range of 0 to `sps_max_sub_layers_minus1 - 1`, inclusive, due to `sps_sub_layer_ordering_info_present_flag` being equal to 0, it is inferred to be equal to `sps_max_dec_pic_buffering_minus1[ sps_max_sub_layers_minus1 ]`.
- [0734] `sps_max_num_reorder_pics[ i ]` indicates the maximum allowed number of pictures that can precede any picture in the CVS in decoding order and follow that picture in output order when `HighestTid` is equal to `i`. The value of `sps_max_num_reorder_pics[ i ]` ranges from 0 to `sps_max_dec_pic_buffering_minus1[ i ]`, inclusive. When `i` is greater than 0, `sps_max_num_reorder_pics[ i ]` may be greater than or equal to `sps_max_num_reorder_pics[ i - 1 ]`. When `sps_max_num_reorder_pics[ i ]` is not present for `i` in the range of 0 to `sps_max_sub_layers_minus1 - 1`, inclusive, due to `sps_sub_layer_ordering_info_present_flag` being equal to 0, it is inferred to be equal to `sps_max_num_reorder_pics[ sps_max_sub_layers_minus1 ]`.
- [0735] `sps_max_latency_increase_plus1[ i ]` not equal to 0 may be used to compute the value of `SpsMaxLatencyPictures[ i ]`, which specifies the maximum number of pictures that can precede any picture in the CVS in output order and follow that picture in decoding order when `HighestTid` is equal to `i`. When `sps_max_latency_increase_plus1[ i ]` is not equal to 0, the value of `SpsMaxLatencyPictures[ i ]` is specified as `SpsMaxLatencyPictures[ i ] = sps_max_num_reorder_pics[ i ] + sps_max_latency_increase_plus1[ i ] - 1`. When `sps_max_latency_increase_plus1[ i ]` is equal to 0, no corresponding limit is expressed.
- [0736] The value of `sps_max_latency_increase_plus1[ i ]` ranges from 0 to  $2^{32} - 2$ , inclusive. When `sps_max_latency_increase_plus1[ i ]` is not present for `i` in the range of 0 to `sps_max_sub_layers_minus1 - 1`, inclusive, due to `sps_sub_layer_ordering_info_present_flag` being equal to 0, it is inferred to be equal to `sps_max_latency_increase_plus1[ sps_max_sub_layers_minus1 ]`.
- [0737]

[Table 44]

seq_parameter_set_rbsp() {
...
sps_max_sub_layers_minus1
...
pic_width_in_luma_samples
pic_height_in_luma_samples
...
for( i = ( sps_sub_layer_ordering_info_present_flag ? 0 :
sps_max_sub_layers_minus1 );
I<=sps_max_sub_layers_minus1; i++ ) {
sps_max_dec_pic_buffering_minus1[i]
sps_max_num_reorder_pics[i]
sps_max_latency_increase_plus1[i]
}
...
}

Table (9)

[0738] Further, JCTVC-L1003 describes the HEVC standard. For example, detail regarding pic\_out\_flag and no\_output\_of\_prior\_pics\_flag is provided in Table (10) below:

[0739] [Table 45]

GENERAL SLICE SEGMENT HEADER SYNTAX
<pre> slice_segment_header() {   first_slice_segment_in_pic_flag   if( nal_unit_type &gt;= BLA_W_LP &amp;&amp; nal_unit_type &lt;= RSV_IRAP_VCL23 )     no_output_of_prior_pics_flag   slice_pic_parameter_set_id   if( !first_slice_segment_in_pic_flag ) {     if( dependent_slice_segments_enabled_flag )       dependent_slice_segment_flag     slice_segment_address   }   if( !dependent_slice_segment_flag ) {     for( i = 0; i &lt; num_extra_slice_header_bits; i++ )       slice_reserved_flag[i]     slice_type     if( output_flag_present_flag )       pic_output_flag     ...     ...   } </pre>

Table (10)

[0740] In Table (10), no\_output\_of\_prior\_pics\_flag affects the output of previously-decoded pictures in the decoded picture buffer (DPB) after the decoding of an IDR or a BLA picture that is not the first picture in the bitstream as specified in Annex C of JCTVC-L1003.

- [0741] output\_flag\_present\_flag equal to 1 indicates that the pic\_output\_flag syntax element is present in the associated slice headers. output\_flag\_present\_flag equal to 0 indicates that the pic\_output\_flag syntax element is not present in the associated slice headers. pic\_output\_flag affects the decoded picture output and removal processes as specified in Annex C of JCTVC-L1003. When pic\_output\_flag is not present, it is inferred to be equal to 1.
- [0742] For the general decoding process (as found in 8.1 of JCTVC-L1003), PicOutputFlag is set as follows:
- [0743] - If the current picture is a RASL picture and NoRaslOutputFlag of the associated IRAP picture is equal to 1, PicOutputFlag may be set equal to 0.
- [0744] - Otherwise, PicOutputFlag may be set equal to pic\_output\_flag.
- [0745] Additionally, during the general decoding process for generating unavailable reference pictures (as found in section 8.3.3.1 of JCTVC-L1003), the value of PicOutputFlag for the generated picture may be set equal to 0 under certain conditions.
- [0746] When the current picture is an IRAP picture, the following applies:
- [0747] - If the current picture is an IDR picture, a BLA picture, the first picture in the bitstream in decoding order or the first picture that follows an end of sequence NAL unit in decoding order, the variable NoRaslOutputFlag may be set equal to 1.
- [0748] - Otherwise, if some external means not specified in JCTVC-L1003 is available to set the variable HandleCraAsBlaFlag to a value for the current picture, the variable HandleCraAsBlaFlag may be set equal to the value provided by the external means and the variable NoRaslOutputFlag may be set equal to HandleCraAsBlaFlag.
- [0749] - Otherwise, the variable HandleCraAsBlaFlag may be set equal to 0 and the variable NoRaslOutputFlag may be set equal to 0.
- [0750] As describe above, access unit (AU) refers to a set of network abstraction layer (NAL) units that are associated with each other according to a specified classification rule, that are consecutive in decoding order, and that include the video coding layer (VCL) NAL units of all coded pictures associated with the same output time and their associated non-VCL NAL units. The base layer is a layer in which all VCL NAL units have a nuh\_layer\_id equal to 0. A coded picture is a coded representation of a picture that includes VCL NAL units with a particular value of nuh\_layer\_id and that includes all the coding tree units of the picture. In some cases, a coded picture may be called a layer component. Addition details about steps being picture based or access unit (AU) based are given in relation to Figures 69 and 70 below.
- [0751] In some configurations, bitstream conformance constraints for pic\_output\_flag and/or no\_output\_of\_prior\_pics\_flag may be used for coded pictures in an access unit (AU). In addition, three new access unit output flags, AU output flag (e.g., AuOutputFlag), AU no RASL output flag (e.g., AuNoRaslOutputFlag) and AU no output

of prior pictures flag (e.g., AuNoOutputOfPriorPicsFlag), may be derived for an AU based on the value of various syntax elements and NAL unit types for the coded pictures in the AU. In some configurations, output and removal of pictures may be based on these three flags (e.g., AuOutputFlag, AuNoRaslOutputFlag and AuNoOutputOfPriorPicsFlag) for SHVC and multi-view HEVC.

- [0752] For example, bitstream conformance constraints for HEVC-extensions may be followed as described in the systems and methods herein. In particular, bitstream conformance constraints may be applied to the SHVC bitstream. In addition, bitstream conformance constraints may be applied for the MV-HEVC bitstream.
- [0753] In one configuration, when present, the value of the slice segment header syntax elements `pic_output_flag` may be required to be the same in all slice segment headers of coded pictures in an access unit (AU). In another configuration, the value of the slice segment header syntax elements `pic_output_flag`, when present, may be the same in all slice segment headers of coded pictures in an access unit (AU) when the coded pictures have the same NAL unit type.
- [0754] In one configuration, when present, the value of the slice segment header syntax elements `pic_output_flag` for the slice segments with `nuh_layer_id` equal to the `nuh_layer_id` value of a target layer may be the same in all slice segment headers of such coded pictures in an access unit (AU). In another configuration, when present, the value of the slice segment header syntax elements `pic_output_flag` for the slice segments with `nuh_layer_id` not equal to the `nuh_layer_id` value of a target layer may be 0 in all slice segment headers of such coded pictures in an access unit (AU).
- [0755] In one configuration, a target layer may be a layer which belongs to a layer set or a target layer set or an output layer set as defined in JCTVC-L1003, JCTVC-M1008, or JCT3V-D1004. In another configuration, a target layer may be a layer which is intended to be decoded. In yet another configuration, a target layer may be a layer which is intended to be decoded and outputted (displayed or otherwise sent for output).
- [0756] In some configurations, when present, the value of the slice segment header syntax elements `no_output_of_prior_pics_flag` may be the same in all slice segment headers of coded pictures in an access unit (AU). In other configurations, when present, the value of the slice segment header syntax elements `no_output_of_prior_pics_flag` may be the same in all slice segment headers of coded pictures in an access unit (AU) when the coded pictures have the same NAL unit type. In one configuration, when present, the value of the slice segment header syntax elements `no_output_of_prior_pics_flag` for the slice segments with `nuh_layer_id` equal to the `nuh_layer_id` value of a target layer may be the same in all slice segment headers of such coded pictures in an access unit (AU).
- [0757] In some configuration, the syntax elements `pic_output_flag` and/or

no\_output\_of\_prior\_pics\_flag may not be signaled when nuh\_layer\_id > 0. In this case, the values for layers with nuh\_layer\_id > 0 may be inferred to be equal to their signaled values for nuh\_layer\_id equal to 0.

- [0758] In some configurations, additional flags, such as AuOutputFlag and AuNoRaslOutputFlag may be employed. The flags AuOutputFlag and AuNoRaslOutputFlag may be derived according to a number of approaches. In one approach or configuration, the two flags, AuOutputFlag and AuNoRaslOutputFlag may be derived and used for DPB operation. AuOutputFlag may be set equal to 1 if PicOutputFlag is equal to 1 for all the pictures in the AU. Otherwise AuOutputFlag may be set equal to 0.
- [0759] In another approach, AuOutputFlag may be set equal to 1 if PicOutputFlag is equal to 1 for at least one picture in the AU. Otherwise AuOutputFlag may be set equal to 0. Thus, in this case, the AuOutputFlag may be set equal to 0 if PicOutputFlag is equal to 0 for all the pictures in an AU.
- [0760] In yet another approach, AuOutputFlag may be set equal to 1 if PicOutputFlag is equal to 1 for pictures belonging to all target output layers in the AU. Otherwise AuOutputFlag may be set equal to 0.
- [0761] In still yet another approach, AuOutputFlag may be set equal to 1 if PicOutputFlag is equal to 1 for picture belonging to at least one target output layers in the AU. Otherwise AuOutputFlag may be set equal to 0.
- [0762] In one approach, AuNoRaslOutputFlag may be set equal to 1 if NoRaslOutputFlag is equal to 1 for all the pictures in the AU. Otherwise AuNoRaslOutputFlag may be set equal to 0.
- [0763] In another approach, AuNoRaslOutputFlag may be set equal to 1 if NoRaslOutputFlag is equal to 1 for at least one picture in the AU. Otherwise AuNoRaslOutputFlag may be set equal to 0. Thus, in this case, the AuNoRaslOutputFlag may be set equal to 0 if NoRaslOutputFlag is equal to 0 for all the pictures in an AU.
- [0764] In yet another approach, AuNoRaslOutputFlag may be set equal to 1 if NoRaslOutputFlag is equal to 1 for pictures belonging to all target output layers in the AU. Otherwise AuNoRaslOutputFlag may be set equal to 0.
- [0765] In still yet another approach, AuNoRaslOutputFlag may be set equal to 1 if NoRaslOutputFlag is equal to 1 for picture belonging to at least one target output layers in the AU. Otherwise AuNoRaslOutputFlag may be set equal to 0.
- [0766] In some of the above approaches and configurations, the DPB operation may use the AuOutputFlag in place of the PicOutputFlag. Additionally, the DPB operation may use AuNoRaslOutputFlag in place of the NoRaslOutputFlag.
- [0767] Examples showing the use of AU output flags, such as AuOutputFlag and AuNoRaslOutputFlag, according to the present systems and methods, are provided below in Listing (1A) and Listing (2A) below. Also, as described below in Listing (1), Listing

(1A), Listing (2) and Listing (2A), the AU output flag `AuNoOutputOfPriorPicsFlag` may be derived and used for DPB operation.

[0768] The present systems and methods may be implemented by changes to standards documents. Listing (1) below provides the sections of JCTVC-L1003 that would be changed to accommodate the present systems and methods.

### Listing 1

[0769] C.3 Operation of the decoded picture buffer (DPB)

#### C.3.1 General

The specifications in this subclause apply independently to each set of DPB parameters selected as specified in subclause C.1. The DPB operates separately or independently for each layer. Thus the following steps take place separately for each decoded picture with a particular value of `nuh_layer_id`.

The decoded picture buffer contains picture storage buffers. Each layer consists of its own set of picture storage buffers. Thus picture storage buffers of each layer are associated with the `nuh_layer_id` value of the layer. Each of the picture storage buffers may contain a decoded picture that is marked as “used for reference” or is held for future output. The processes specified in subclauses C.3.2, C.3.3 and C.3.4 are sequentially applied as specified below.

#### C.3.2 Removal of pictures from the DPB

The removal of pictures from the DPB before decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously at the CPB removal time of the first decoding unit of the current picture belonging to access unit *n* (containing the current picture) and proceeds as follows:

- The decoding process for RPS as specified in subclause 8.3.2 is invoked.
- When the current picture is an IRAP picture with `NoRaslOutputFlag` equal to 1 that is not picture 0, the following ordered steps are applied:
  1. The variable `NoOutputOfPriorPicsFlag` is derived for the decoder under test as follows:
    - If the current picture is a CRA picture, `NoOutputOfPriorPicsFlag` is set equal to 1 (regardless of the value of `no_output_of_prior_pics_flag`).
    - Otherwise, if the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]` derived from the active SPS corresponding to the `nuh_layer_id` value of the current picture is different from the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, respectively, derived from the SPS active for the preceding picture with the

nuh\_layer\_id value equal to the nuh\_layer\_id value of the current picture, NoOutputOfPriorPicsFlag may (but should not) be set to 1 by the decoder under test, regardless of the value of no\_output\_of\_prior\_pics\_flag.

NOTE – Although setting NoOutputOfPriorPicsFlag equal to no\_output\_of\_prior\_pics\_flag is preferred under these conditions, the decoder under test is allowed to set NoOutputOfPriorPicsFlag to 1 in this case.

- Otherwise, NoOutputOfPriorPicsFlag is set equal to no\_output\_of\_prior\_pics\_flag.
- 2. The value of NoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers corresponding to the nuh\_layer\_id value of the current picture in the DPB are emptied without output of the pictures they contain, and the DPB fullness for the nuh\_layer\_id value of the current picture is set equal to 0.

In one embodiment the value of NoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers corresponding to all the nuh\_layer\_id values in the DPB are emptied without output of the pictures they contain, and the DPB fullness for all nuh\_layer\_id values is set equal to 0.

In one embodiment the value of NoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers PSB[currLayerId] with currLayerId equal to nuh\_layer\_id value of the current picture in the DPB are emptied without output of the pictures they contain, and the DPB fullness DPBFullness[currLayerId] for the nuh\_layer\_id value currLayerId of the current picture is set equal to 0.

In one embodiment the value of NoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers PSB[nuh\_layer\_id] for all the nuh\_layer\_id values in the DPB are emptied without output of the pictures they contain, and the DPB fullness DPBFullness[nuh\_layer\_id] for all nuh\_layer\_id values is set equal to 0.



- When both of the following conditions are true for any pictures  $k$  in the picture storage buffer corresponding to the  $\text{nuh\_layer\_id}$  value of the current picture in the DPB, all such pictures  $k$  in the DPB are removed from the DPB:
  - picture  $k$  is marked as “unused for reference”
  - picture  $k$  has  $\text{PicOutputFlag}$  equal to 0 or its DPB output time is less than or equal to the CPB removal time of the first decoding unit (denoted as decoding unit  $m$ ) of the current picture  $n$ ; i.e.  $\text{DpbOutputTime}[k]$  is less than or equal to  $\text{CpbRemovalTime}(m)$
  - For each picture that is removed from the DPB, the DPB fullness is decremented by one.
  - In one embodiment for each picture  $k$  having  $\text{nuh\_layer\_id}$  value  $\text{nuhLayerIdk}$  that is removed from the DPB, the DPB fullness  $\text{DPBfullness}[\text{nuhLayerIdk}]$  is decremented by one.

### C.3.3 Picture output

The processes specified in this subclause happen instantaneously at the CPB removal time of access unit  $n$ ,  $\text{AuCpbRemovalTime}[n]$ .

When picture  $n$  has  $\text{PicOutputFlag}$  equal to 1, its DPB output time  $\text{DpbOutputTime}[n]$  is derived as follows, where the variable  $\text{firstPicInBufferingPeriodFlag}$  is equal to 1 if access unit  $n$  is the first access unit of a buffering period and 0 otherwise:

```

if( !SubPicHrdFlag ) {
    DpbOutputTime[ n ] = AuCpbRemovalTime[ n ] + ClockTick * picDpbOutputDelay (C-16)
    if( firstPicInBufferingPeriodFlag )
        DpbOutputTime[ n ] -= ClockTick * DpbDelayOffset
    } else
        DpbOutputTime[ n ] = AuCpbRemovalTime[ n ] + ClockSubTick *
picSptDpbOutputDuDelay

```

where  $\text{picDpbOutputDelay}$  is the value of  $\text{pic\_dpb\_output\_delay}$  in the picture timing SEI message associated with access unit  $n$ , and  $\text{picSptDpbOutputDuDelay}$  is the value of  $\text{pic\_spt\_dpb\_output\_du\_delay}$ , when present, in the decoding unit information SEI messages associated with access unit  $n$ , or the value of  $\text{pic\_dpb\_output\_du\_delay}$  in the picture timing SEI message associated with access unit  $n$  when there is no decoding unit information SEI message

associated with access unit n or no decoding unit information SEI message associated with access unit n has `pic_spt_dpb_output_du_delay` present.

NOTE – When the syntax element `pic_spt_dpb_output_du_delay` is not present in any decoding unit information SEI message associated with access unit n, the value is inferred to be equal to `pic_dpb_output_du_delay` in the picture timing SEI message associated with access unit n.

The output of the current picture if its `nuh_layer_id` belongs to the layer in the `TargetDecLayerIdList` is specified as follows:

- If `PicOutputFlag` is equal to 1 and `DpbOutputTime[ n ]` is equal to `AuCpbRemovalTime[ n ]`, the current picture is output.
- Otherwise, if `PicOutputFlag` is equal to 0, the current picture is not output, but will be stored in the picture storage buffer corresponding to the `nuh_layer_id` value of the current picture in the DPB as specified in subclause C.3.4.

In one embodiment: Otherwise, if `PicOutputFlag` is equal to 0, the current picture is not output, but will be stored in the picture storage buffer `PSB[currLayerId]` corresponding to the `nuh_layer_id` value `currLayerId` of the current picture in the DPB as specified in subclause C.3.4.

- Otherwise ( `PicOutputFlag` is equal to 1 and `DpbOutputTime[ n ]` is greater than `AuCpbRemovalTime[ n ]` ), the current picture is output later and will be stored in the picture storage buffer corresponding to the `nuh_layer_id` value of the current picture in the DPB (as specified in subclause C.3.4) and is output at time `DpbOutputTime[ n ]` unless indicated not to be output by the decoding or inference of `no_output_of_prior_pics_flag` equal to 1 at a time that precedes `DpbOutputTime[ n ]`.

In another embodiment the above steps are specified for:

The output of the current picture if its `nuh_layer_id` belongs to a layer which belongs to the output layer set corresponding to the (current) operation point.

In yet another embodiment the above steps are specified for:

The output of current picture (without checking if it belongs to the `TargetDecLayerIdList` or the output layer set for the current operation point).

When output, the picture is cropped, using the conformance cropping window specified in the active SPS for the picture.

When picture  $n$  is a picture that is output and is not the last picture of the bitstream that is output, the value of the variable  $DpbOutputInterval[n]$  is derived as follows:

$$DpbOutputInterval[n] = DpbOutputTime[nextPicInOutOrder] - DpbOutputTime[n] \quad (C-17)$$

where  $nextPicInOutOrder$  is the picture that follows picture  $n$  in output order and has  $PicOutputFlag$  equal to 1.

#### C.3.4 Current decoded picture marking and storage

The process specified in this subclause happens instantaneously at the CPB removal time of access unit  $n$ ,  $CpbRemovalTime[n]$ .

The current decoded picture is stored in the DPB in an empty picture storage buffer corresponding to the  $nuh\_layer\_id$  value of the current picture, the DPB fullness for the  $nuh\_layer\_id$  value of the current picture is incremented by one, and the current picture is marked as “used for short-term reference.”

In one embodiment:

The current decoded picture with  $nuh\_layer\_id$  equal to  $currLayerId$  is stored in the DPB in an empty picture storage buffer  $PSB[currLayerId]$  with  $currLayerId$  equal to  $nuh\_layer\_id$  value of the current picture in the DPB, the DPB fullness for the  $nuh\_layer\_id$  value of the current picture  $DPBFullness[currLayerId]$  is incremented by one, and the current picture is marked as “used for short-term reference.”

#### C.4 Bitstream conformance

The specifications in subclause C.4 apply.

#### C.5 Decoder conformance

##### F.8.1.1 C.5.1 General

The specifications in subclause C.5.1 apply.

#### Listing (1)

[0770] As used in Listing (1) above, PSB refers to a picture storage buffer.  $DPBFullness$  refers to a variable used to describe the fullness of decoded picture buffer (DPB) 116.

[0771] Listing (1A) below provides an alternative approach to section C.3.2 of Listing (1) according to accommodate the present systems and methods. In some configurations, Listing (1A) may only represent changes to section C.3.2 in JCTVC-L1003. Listing (1A) may use the flags  $AuNoOutputOfPriorPicsFlag$  and  $AuNoRaslOutputFlag$  defined above.

#### Listing 1A

[0772]

### C.3.2 Removal of pictures from DPB

The removal of pictures from the DPB before decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously at the CPB removal time of the first decoding unit of the current picture belonging to access unit *n* (containing the current picture) and proceeds as follows:

- The decoding process for RPS as specified in subclause 8.3.2 is invoked.
- When the current picture is an IRAP picture with `AuNoRaslOutputFlag` equal to 1 that is not picture 0, the following ordered steps are applied:
- In another configuration, when the current picture is an IRAP picture with `NoRaslOutputFlag` equal to 1 that is not picture 0, the following ordered steps are applied:
  1. The variable `AuNoOutputOfPriorPicsFlag` is derived for the decoder under test as follows:
    - If the current picture is a CRA picture, `AuNoOutputOfPriorPicsFlag` is set equal to 1 (regardless of the value of `no_output_of_prior_pics_flag` for the current picture or other pictures in the AU).
    - Otherwise, if the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]` derived from the active SPS is different from the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, respectively, derived from the SPS active for the preceding picture, `AuNoOutputOfPriorPicsFlag` may (but should not) be set to 1 by the decoder under test, regardless of the value of `no_output_of_prior_pics_flag`.

NOTE – Although setting `AuNoOutputOfPriorPicsFlag` equal to `no_output_of_prior_pics_flag` is preferred under these conditions, the decoder under test is allowed to set `AuNoOutputOfPriorPicsFlag` to 1 in this case.

- Otherwise, AuNoOutputOfPriorPicsFlag is set based on the value of no\_output\_of\_prior\_pics\_flag for the current picture and other pictures in the AU as follows:
    - AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for at least one picture in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0. Thus in this case the AuNoOutputOfPriorPicsFlag is set equal to 0 if no\_output\_of\_prior\_pics\_flag is equal to 0 for all the pictures in an AU.
    - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for the current picture. Otherwise AuNoOutputOfPriorPicsFlag is left unchanged.
    - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for all the pictures in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0.
    - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for all the pictures belonging to the target output layers in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0.
    - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for at least one picture belonging to the target output layers in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0.
2. The value of AuNoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of AuNoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers corresponding to all the nuh\_layer\_id values in the DPB are emptied without output of the pictures they contain, and the DPB fullness for all the nuh\_layer\_id values is set equal to 0.
- In another embodiment, the value of AuNoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD, such that when the value of AuNoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers corresponding to the nuh\_layer\_id value of the current picture in the DPB are emptied without output of the pictures they contain, and the DPB fullness for the nuh\_layer\_id value of the current picture is set equal to 0.

Listing (1A)

[0773] Listing (2) below provides the sections of JCTVC-L1008 that would be changed to accommodate the present systems and methods.

**Listing 2**

[0774]

### F.13 Hypothetical reference decoder

#### F.13.1 General

The specifications in subclause C.1 apply.

#### F.13.2 Operation of coded picture buffer (CPB)

The specifications in subclause C.2 apply.

#### F.13.3 Operation of the decoded picture buffer (DPB)

The specifications in subclause C.3 apply separately for each set of decoded pictures with a particular value of `nuh_layer_id`.

`PicOutputFlag` for pictures that are not included in a target output layer is set equal to 0.

Decoded pictures with the same DPB output time and with `PicOutputFlag` equal to 1 are output in ascending order of `nuh_layer_id` values of these decoded pictures.

#### F.13.5 Decoder conformance

##### F.13.5.1 General

The specifications in subclause C.5.1 apply.

##### F.13.5.2 Operation of the output order DPB

###### F.13.5.2.1 General

The decoded picture buffer contains picture storage buffers. Each layer consists of its own set of picture storage buffers. Thus picture storage buffers of each layer are associated with the `nuh_layer_id` value of the layer. The number of picture storage buffers for `nuh_layer_id` equal to 0 is derived from the active SPS of the layer with `nuh_layer_id` equal to 0. The number of picture storage buffers for each non-zero `nuh_layer_id` value is derived from the active layer SPS for that non-zero `nuh_layer_id` value. Each of the picture storage buffers contains a decoded picture that is marked as “used for reference” or is held for future output. The process for output and removal of pictures from the DPB as specified in subclause F.13.5.2.2 is invoked, followed by

the invocation of the process for picture decoding, marking, additional bumping, and storage as specified in subclause F.13.5.2.3. The “bumping” process is specified in subclause F.13.5.2.4 and is invoked as specified in subclauses F.13.5.2.2 and F.13.5.2.3.

#### F.13.5.2.2 Output and removal of pictures from the DPB

The output and removal of pictures from the DPB before the decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously when the first decoding unit of the current picture belonging to the access unit containing the current picture is removed from the CPB and proceeds as follows.

The decoding process for RPS as specified in subclause 8.3.2 is invoked.

- If the current picture is an IRAP picture with NoRaslOutputFlag equal to 1 and with nuh\_layer\_id equal to 0 that is not picture 0, the following ordered steps are applied:
    1. The variable NoOutputOfPriorPicsFlag is derived for the decoder under test as follows:
      - If the current picture is a CRA picture, NoOutputOfPriorPicsFlag is set equal to 1 (regardless of the value of no\_output\_of\_prior\_pics\_flag).
      - Otherwise, if the value of pic\_width\_in\_luma\_samples, pic\_height\_in\_luma\_samples, or sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ] derived from the active SPS corresponding to the nuh\_layer\_id value of the current picture is different from the value of pic\_width\_in\_luma\_samples, pic\_height\_in\_luma\_samples, or sps\_max\_dec\_pic\_buffering\_minus1[ HighestTid ], respectively, derived from the SPS active for the preceding picture with the nuh\_layer\_id value equal to the nuh\_layer\_id value of the current picture, NoOutputOfPriorPicsFlag may (but should not) be set to 1 by the decoder under test, regardless of the value of no\_output\_of\_prior\_pics\_flag.
- NOTE – Although setting NoOutputOfPriorPicsFlag equal to no\_output\_of\_prior\_pics\_flag is preferred under these conditions, the decoder under test is allowed to set NoOutputOfPriorPicsFlag to 1 in this case.
- Otherwise, NoOutputOfPriorPicsFlag is set equal to no\_output\_of\_prior\_pics\_flag.
  - 2. The value of NoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD as follows:
    - If NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers corresponding to the nuh\_layer\_id value of the current picture in the DPB are emptied without

output of the pictures they contain, and the DPB fullness for the nuh\_layer\_id value of the current picture is set equal to 0.

In one embodiment if NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers corresponding to all the nuh\_layer\_id values in the DPB are emptied without output of the pictures they contain, and the DPB fullness for all nuh\_layer\_id values is set equal to 0.

In one embodiment if NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers PSB[currLayerId] corresponding to the nuh\_layer\_id value currLayerId of the current picture in the DPB are emptied without output of the pictures they contain, and the DPB fullness DPBFullness[currLayerId] is set equal to 0.

In one embodiment if NoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers PSB[nuh\_layer\_id] for all the nuh\_layer\_id values in the DPB are emptied without output of the pictures they contain, and the DPB fullness DPBFullness[nuh\_layer\_id] for all nuh\_layer\_id values is set equal to 0.

- Otherwise (NoOutputOfPriorPicsFlag is equal to 0), all picture storage buffers containing a picture that is marked as “not needed for output” and “unused for reference” are emptied (without output), and all non-empty picture storage buffers in the DPB corresponding to the nuh\_layer\_id value of the current picture are emptied by repeatedly invoking the “bumping” process specified in subclause F.13.5.2.4, and the DPB fullness for the nuh\_layer\_id value of the current picture is set equal to 0.

In another embodiment when NoOutputOfPriorPicsFlag is equal to 0, all picture storage buffers corresponding to all the nuh\_layer\_id values containing a picture that is marked as “not needed for output” and “unused for reference” are emptied (without output), and all non-empty picture storage buffers corresponding to all the nuh\_layer\_id values in the DPB are emptied by repeatedly invoking the “bumping” process specified in subclause F.13.5.2.4, and the DPB fullness for all nuh\_layer\_id values is set equal to 0.

- In another embodiment Otherwise (NoOutputOfPriorPicsFlag is equal to 0), all picture storage buffers containing a picture that is marked as “not needed for output” and “unused for reference” are emptied (without output), and all non-empty picture



storage buffers PSB[currLayerId] in the DPB corresponding to the nuh\_layer\_id value currLayerId of the current picture are emptied by repeatedly invoking the “bumping” process specified in subclause F.13.5.2.4, and the DPB fullness DPBFullness[currLayerId] for the nuh\_layer\_id value of the current picture is set equal to 0.

- In another embodiment when NoOutputOfPriorPicsFlag is equal to 0, all picture storage buffers PSB[nuh\_layer\_id] for all the nuh\_layer\_id values in the DPB containing a picture that is marked as “not needed for output” and “unused for reference” are emptied (without output), and all non-empty picture storage buffers PSB[nuh\_layer\_id] corresponding to all the nuh\_layer\_id values in the DPB are emptied by repeatedly invoking the “bumping” process specified in subclause F.13.5.2.4, and the DPB fullness DPBFullness[nuh\_layer\_id] for all nuh\_layer\_id values is set equal to 0.
- Otherwise (the current picture is not an IRAP picture with NoRaslOutputFlag equal to 1 and with nuh\_layer\_id equal to 0), all picture storage buffers corresponding to the nuh\_layer\_id value of the current picture containing a picture which are marked as “not needed for output” and “unused for reference” are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness corresponding to the nuh\_layer\_id value of the current decoded picture is decremented by one. The variable currLayerId is set equal to nuh\_layer\_id of the current decoded picture and when one or more of the following conditions are true, the “bumping” process specified in subclause F.13.5.2.4 is invoked repeatedly while further decrementing the DPB fullness by one for each additional picture storage buffer corresponding to the nuh\_layer\_id value of the current decoded picture that is emptied, until none of the following conditions are true:
  - The number of pictures with nuh\_layer\_id equal to currLayerId in the DPB that are marked as “needed for output” is greater than sps\_max\_num\_reorder\_pics[ HighestTid ] from the active SPS (when currLayerId is equal to 0) or from the active layer SPS for the value of currLayerId.
  - sps\_max\_latency\_increase\_plus1[ HighestTid ] of the active SPS (when currLayerId is equal to 0) or the active layer SPS for the value of currLayerId is not equal to 0 and there is at least one picture with nuh\_layer\_id equal to currLayerId in the DPB that is marked

as “needed for output” for which the associated variable `PicLatencyCount[ currLayerId ]` is greater than or equal to `SpsMaxLatencyPictures[ HighestTid ]` derived from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.

- The number of pictures with `nuh_layer_id` equal to `currLayerId` in the DPB is greater than or equal to `sps_max_dec_pic_buffering_minus1[ HighestTid ] + 1` from the active SPS (when `currLayerId` is equal to 0) or from the active layer SPS for the value of `currLayerId`.

In another embodiment : Otherwise (the current picture is not an IRAP picture with `NoRaslOutputFlag` equal to 1 and with `nuh_layer_id` equal to 0), all picture storage buffers corresponding to all the `nuh_layer_id` values containing a picture which are marked as “not needed for output” and “unused for reference” are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness corresponding to the `nuh_layer_id` value of the picture is decremented by one.

In another embodiment: Otherwise (the current picture is not an IRAP picture with `NoRaslOutputFlag` equal to 1 and with `nuh_layer_id` equal to 0), all picture storage buffers `PSB[currLayerId]` in the DPB corresponding to the `nuh_layer_id` value `currLayerId` of the current picture containing a picture which are marked as “not needed for output” and “unused for reference” are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness `DPBfullness[currLayerId]` corresponding to the `nuh_layer_id` value of the current decoded picture is decremented by one.

In another embodiment: Otherwise (the current picture is not an IRAP picture with `NoRaslOutputFlag` equal to 1 and with `nuh_layer_id` equal to 0), all picture storage buffers `PSB[nuh_layer_id]` for all the `nuh_layer_id` values in the DPB containing a picture which are marked as “not needed for output” and “unused for reference” are emptied (without output). For each picture storage buffer that is emptied, the DPB fullness `DPBFullness[nuh_layer_id]` corresponding to the `nuh_layer_id` value of the picture emptied is decremented by one.

#### F.13.5.2.3 Picture decoding, marking, additional bumping, and storage

The processes specified in this subclause happen instantaneously when the last decoding unit of access unit `n` containing the current picture is removed from the CPB.

The variable currLayerId is set equal to nuh\_layer\_id of the current decoded picture.

For each picture in the DPB that is marked as “needed for output” and that has a nuh\_layer\_id value equal to currLayerId, the associated variable PicLatencyCount[ currLayerId ] is set equal to PicLatencyCount[ currLayerId ] + 1.

The current picture is considered as decoded after the last decoding unit of the picture is decoded. The current decoded picture is stored in an empty picture storage buffer corresponding to currLayerId (the nuh\_layer\_id value of the current picture) in the DPB, and the following applies:

- If the current decoded picture has PicOutputFlag equal to 1, it is marked as “needed for output” and its associated variable PicLatencyCount[ currLayerId ] is set equal to 0.
- Otherwise (the current decoded picture has PicOutputFlag equal to 0), it is marked as “not needed for output.”

In one embodiment, the current decoded picture is stored in an empty picture storage buffer corresponding to currLayerId (the nuh\_layer\_id value of the current picture) in the DPB, and the following applies:

- If the current decoded picture has PicOutputFlag equal to 1, it is marked as “needed for output” and its associated variable PicLatencyCount[ currLayerId ] is set equal to 0. All picture storage buffers containing a picture that has the same picture order count value (PicOrderCntVal ) as the current decoded picture are marked as “needed for output.”
- Otherwise (the current decoded picture has PicOutputFlag equal to 0), it is marked as “not needed for output.”

In one embodiment, the current decoded picture is stored in an empty picture storage buffer PSB[currLayerId] corresponding to currLayerId (the nuh\_layer\_id value of the current picture) in the DPB, and the DPB fullness for the nuh\_layer\_id value of the current picture DPBFullness[currLayerId] is incremented by one, and the following applies

The current decoded picture is marked as “used for short-term reference.”

When one or more of the following conditions are true, the “bumping” process specified in subclause F.13.5.2.4 is invoked repeatedly until none of the following conditions are true.

- The number of pictures with nuh\_layer\_id equal to currLayerId in the DPB that are marked as “needed for output” is greater than sps\_max\_num\_reorder\_pics[ HighestTid ] from the

active SPS (when currLayerId is equal to 0) or from the active layer SPS for the value of currLayerId.

- sps\_max\_latency\_increase\_plus1[ HighestTid ] is not equal to 0 and there is at least one picture with nuh\_layer\_id equal to currLayerId in the DPB that is marked as “needed for output” for which the associated variable PicLatencyCount[ currLayerId ] that is greater than or equal to SpsMaxLatencyPictures[ HighestTid ] derived from the active SPS (when currLayerId is equal to 0) or from the active layer SPS for the value of currLayerId..

#### F.13.5.2.4 “Bumping” process

The “bumping” process consists of the following ordered steps:

1. The pictures that are first for output are selected as the ones having the smallest value of PicOrderCntVal of all pictures in the DPB marked as “needed for output.”
2. These pictures are cropped, using the conformance cropping window specified in the active SPS for the picture with nuh\_layer\_id equal to 0 or in the active layer SPS for a nuh\_layer\_id value equal to that of the picture, the cropped pictures are output in ascending order of nuh\_layer\_id, and the pictures are marked as “not needed for output.”
3. Each picture storage buffer that contains a picture marked as “unused for reference” and that included one of the pictures that was cropped and output is emptied.

#### Listing (2)

[0775] Listing (2A) below provides an alternative approach to section F.13.5.2.2 of Listing (2) according to accommodate the present systems and methods. In some configurations, Listing (2A) may only represent changes to section F.13.5.2.2 in JCTVC-L1008. Listing (2A) may use the flags AuNoOutputOfPriorPicsFlag and AuNo-RaslOutputFlag defined above.

#### Listing 2A

[0776]

#### F.13.5.2.2 Output and removal of pictures from the DPB

The output and removal of pictures from the DPB before the decoding of the current picture (but after parsing the slice header of the first slice of the current picture) happens instantaneously when the first decoding unit of the access unit containing the current picture is removed from the CPB and proceeds as follows:

The decoding process for RPS as specified in subclause 8.3.2 is invoked.

- If the current picture is an IRAP picture with `AuNoRaslOutputFlag` equal to 1 and with `nuh_layer_id` equal to 0 that is not picture 0, the following ordered steps are applied:
- In another configuration, if the current picture is an IRAP picture with `NoRaslOutputFlag` equal to 1 and with `nuh_layer_id` equal to 0 that is not picture 0, the following ordered steps are applied:
  1. The variable `AuNoOutputOfPriorPicsFlag` is derived for the decoder under test as follows.
    - If current picture has `nuh_layer_id` equal to 0 `AuNoOutputOfPriorPicsFlag` is initialized to 0. Then:
      - If the current picture is a CRA picture, `AuNoOutputOfPriorPicsFlag` is set equal to 1 (regardless of the value of `no_output_of_prior_pics_flag` for the current picture or other pictures in the AU).
      - Otherwise, if the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]` derived from the active SPS is different from the value of `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, or `sps_max_dec_pic_buffering_minus1[ HighestTid ]`, respectively, derived from the SPS active for the preceding picture, `AuNoOutputOfPriorPicsFlag` may (but should not) be set to 1 by the decoder under test, regardless of the value of `no_output_of_prior_pics_flag`.

NOTE – Although setting AuNoOutputOfPriorPicsFlag equal to no\_output\_of\_prior\_pics\_flag is preferred under these conditions, the decoder under test is allowed to set AuNoOutputOfPriorPicsFlag to 1 in this case.

- Otherwise, AuNoOutputOfPriorPicsFlag is set based on the value of no\_output\_of\_prior\_pics\_flag for the current picture and other pictures in the AU as follows:
  - AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for at least one picture in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0. Thus in this case the AuNoOutputOfPriorPicsFlag is set equal to 0 if no\_output\_of\_prior\_pics\_flag is equal to 0 for all the pictures in an AU.
  - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for current picture. Otherwise AuNoOutputOfPriorPicsFlag is left unchanged.
  - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for all the pictures in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0.
  - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for all the pictures belonging to the target output layers in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0.
  - In another embodiment AuNoOutputOfPriorPicsFlag is set equal to 1 if no\_output\_of\_prior\_pics\_flag is equal to 1 for at least one picture belonging to the target output layer in the AU. Otherwise AuNoOutputOfPriorPicsFlag is set equal to 0.
- 2. The value of AuNoOutputOfPriorPicsFlag derived for the decoder under test is applied for the HRD as follows:
  - If AuNoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers in the DPB corresponding to all the nuh\_layer\_id values are emptied without output of the

pictures they include, and the DPB fullness is set equal to 0 for all the nuh\_layer\_id values.

- In another embodiment, if AuNoOutputOfPriorPicsFlag is equal to 1, all picture storage buffers in the DPB corresponding to all the nuh\_layer\_id value of the current picture are emptied without output of the pictures they include, and the DPB fullness is set equal to 0 for the nuh\_layer\_id value of the current picture.

Otherwise (AuNoOutputOfPriorPicsFlag is equal to 0), all picture storage buffers corresponding to all the nuh\_layer\_id values including a picture that is marked as “not needed for output” and “unused for reference” are emptied (without output), and all non-empty picture storage buffers in the DPB are emptied by repeatedly invoking the “bumping” process specified in subclause F.13.5.2.4, and the DPB fullness for all the nuh\_layer\_id values is set equal to 0.

In another embodiment:

Otherwise (AuNoOutputOfPriorPicsFlag is equal to 0), all picture storage buffers corresponding to the nuh\_layer\_id value of the current picture including a picture that is marked as “not needed for output” and “unused for reference” are emptied (without output), and all non-empty picture storage buffers in the DPB are emptied by repeatedly invoking the “bumping” process specified in subclause F.13.5.2.4, and the DPB fullness for the nuh\_layer\_id value of the current picture is set equal to 0.

Listing (2A)

- [0777] In one configuration, for the proposed text in Listing (1), Listing (1A), Listing (2) and Listing (2A) above, all occurrences of “current picture” may be replaced by “current decoded picture”.
- [0778] Listing (3) below provides an additional section for decoding to accommodate the present systems and methods. Listing (3) provides the sections of JCTVC-L1003 (i.e., the HEVC specification version 34) that would be changed to accommodate the present systems and methods.

### **Listing 3**

[0779]

## 8.3.2. Decoding process for reference picture set

When the current picture is an IRAP picture with NoRaslOutputFlag equal to 1, all reference pictures currently in the DPB corresponding to the nuh\_layer\_id value of the current picture (if any) are marked as “unused for reference.”

In another embodiment: When the current picture is an IRAP picture with NoRaslOutputFlag equal to 1, all reference pictures currently in the DPB corresponding to all the nuh\_layer\_id values (if any) are marked as “unused for reference.”

## Listing (3)

[0780] Many different variants may be used for defining hybrid decoded picture buffer (DPB) 116 operations with various steps being picture based or access unit based. Tables (11) -(15) below list multiple variants of the steps and corresponding picture based/access unit (AU) based settings.

[0781] [Table 46]

Step	Picture Based / AU Based
Removal	Picture based
Output	AU Based
Storage	Picture Based
Marking	AU Based for output marking, Picture based for reference marking
Additional Output	AU Based

Table (11)

[0782]



[Table 47]

Step	Picture Based / AU Based
Removal	Picture based
Output	AU based
Storage	AU based
Marking	AU based for output marking, AU based for reference marking
Additional Output	AU based

Table (12)

[0783] [Table 48]

Step	Picture Based / AU Based
Removal	AU based
Output	AU based
Storage	AU based
Marking	AU based for output marking, AU based for reference marking
Additional Output	AU based

Table (13)

[0784] [Table 49]

Step	Picture Based / AU Based
Removal	Picture based
Output	AU based
Storage	Picture based
Marking	Picture based for output marking, Picture based for reference marking
Additional Output	AU based

Table (14)

[0785]

[Table 50]

Step	Picture Based / AU Based
Removal	Picture based
Output	Picture based
Storage	Picture based
Marking	Picture based for output marking, Picture based for reference marking
Additional Output	Picture based

Table (15)

[0786] Figure 66 is a block diagram illustrating one configuration of a decoder 2404. The decoder 2404 may be included in an electronic device 2402. For example, the decoder 2404 may be a high-efficiency video coding (HEVC) decoder. The decoder 2404 and/or one or more of the elements illustrated as included in the decoder 2404 may be implemented in hardware, software or a combination of both. The decoder 2404 may receive a bitstream 2414 (e.g., one or more encoded pictures included in the bitstream 2414) for decoding. In some configurations, the received bitstream 2414 may include received overhead information, such as a received slice header, received PPS, received buffer description information, etc. The encoded pictures included in the bitstream 2414 may include one or more encoded reference pictures and/or one or more other encoded pictures.

[0787] Received symbols (in the one or more encoded pictures included in the bitstream 2414) may be entropy decoded by an entropy decoding module 454, thereby producing a motion information signal 456 and quantized, scaled and/or transformed coefficients 458.

[0788] The motion information signal 456 may be combined with a portion of a reference frame signal 484 from a frame memory 464 at a motion compensation module 460, which may produce an inter-frame prediction signal 468. The quantized, descaled and/or transformed coefficients 458 may be inverse quantized, scaled and inverse transformed by an inverse module 462, thereby producing a decoded residual signal 470. The decoded residual signal 470 may be added to a prediction signal 478 to produce a combined signal 472. The prediction signal 478 may be a signal selected from either the inter-frame prediction signal 468 or an intra-frame prediction signal 476 produced by an intra-frame prediction module 474. In some configurations, this signal selection may be based on (e.g., controlled by) the bitstream 2414.

- [0789] The intra-frame prediction signal 476 may be predicted from previously decoded information from the combined signal 472 (in the current frame, for example). The combined signal 472 may also be filtered by a de-blocking filter 480. The resulting filtered signal 482 may be written to frame memory 464. The resulting filtered signal 482 may include a decoded picture.
- [0790] The frame memory 464 may include a decoded picture buffer (DPB) 2416 as described herein. The decoded picture buffer (DPB) 2416 may be capable of hybrid decoded picture buffer (DPB) 116 operations. The decoded picture buffer (DPB) 2416 may include one or more decoded pictures that may be maintained as short or long term reference frames. The frame memory 464 may also include overhead information corresponding to the decoded pictures. For example, the frame memory 464 may include slice headers, video parameter set (VPS) information, sequence parameter set (SPS) information, picture parameter set (PPS) information, cycle parameters, buffer description information, etc. One or more of these pieces of information may be signaled from an encoder (e.g., encoder 2108, overhead signaling module 2112).
- [0791] Figure 67A is a block diagram illustrating the use of both an enhancement layer and a base layer for video coding with separate decoded picture buffers (DPBs) 516a-b and separate hybrid decoded picture buffer (DPB) operation modules 520a-b for the base layer and the enhancement layer. A first electronic device 502a and a second electronic device 502b are illustrated. The first electronic device 502a may include a video encoder 508 that includes an enhancement layer encoder 526 and a base layer encoder 528. Each of the elements included within the first electronic device 502a (i.e., the enhancement layer encoder 526 and the base layer encoder 528) may be implemented in hardware, software or a combination of both. The first electronic device 502a may obtain an input picture 2506. In some configurations, the input picture 2506 may be captured on the first electronic device 502a using an image sensor, retrieved from memory or received from another electronic device 502.
- [0792] The enhancement layer encoder 526 may encode the input picture 2506 to produce encoded data. For example, the enhancement layer encoder 526 may encode a series of input pictures 2506 (e.g., video). The encoded data may be included in an encoded enhancement layer video bitstream 530. The enhancement layer encoder 526 may generate overhead signaling based on the input picture 2506.
- [0793] The enhancement layer video decoder 534 may include a decoded picture buffer (DPB) 516a and a hybrid decoded picture buffer (DPB) operation module 520a. Likewise, the base layer decoder 536 may include a decoded picture buffer (DPB) 516b and a hybrid decoded picture buffer (DPB) operation module 520b.
- [0794] The base layer encoder 528 may also encode the input picture 2506. In one configuration, the same input picture 2506 used by the enhancement layer encoder 526

may also be used by the base layer encoder 528. In another configuration, a different (but similar) input picture than the input picture 2506 used by the enhancement layer encoder 526 may be used by the base layer encoder 528. For example, for signal-to-noise ratio (SNR) scalability (also referred to as quality scalability), the same input picture 2506 may be used by both the enhancement layer encoder 526 and the base layer encoder 528. As another example, for spatial scalability, a downsampled picture may be used by the base layer encoder 528. In yet another example, for multi-view scalability, a different view picture may be used by the base layer encoder 528. The base layer encoder 528 may produce encoded data included in an encoded base layer video bitstream 532.

- [0795] The encoded enhancement layer video bitstream 530 and the encoded base layer video bitstream 532 may each include encoded data based on the input picture 2506. In one example, the encoded enhancement layer video bitstream 530 and the encoded base layer video bitstream 532 may include encoded picture data. In some configurations, the encoded enhancement layer video bitstream 530 and/or the encoded base layer video bitstream 532 may also include overhead data, such as sequence parameter set (SPS) information, picture parameter set (PPS) information, video parameter set (VPS) information, slice header information, etc.
- [0796] The encoded enhancement layer video bitstream 530 may be provided to the second electronic device 502b. Likewise, the encoded base layer video bitstream 532 may be provided to the second electronic device 502b. The second electronic device 502b may include a video decoder 2504. The video decoder 2504 may include an enhancement layer decoder 534 and a base layer decoder 536. In one configuration, the encoded base layer video bitstream 532 is decoded by the base layer decoder 536 while the encoded enhancement layer video bitstream 530 is decoded by the enhancement layer decoder 534.
- [0797] In one example, the encoded enhancement layer video bitstream 530 and the encoded base layer video bitstream 532 may be transmitted to the second electronic device 502b using a wired or wireless link. In some cases, this may be done over a network, such as the Internet, a Local Area Network (LAN) or other type of network for communicating between devices. It should be noted that in some configurations, the encoders (i.e., the enhancement layer encoder 526 and the base layer encoder 528) and the decoder 2504 (e.g., the base layer decoder 536 and the enhancement layer decoder 534) may be implemented on the same electronic device 502 (i.e., the first electronic device 502a and the second electronic device 502b may be part of a single electronic device 502). In an implementation where the encoders and decoders are implemented on the same electronic device 502, for instance, the encoded enhancement layer video bitstream 530 and the encoded base layer video bitstream 532 may be made available to the

video decoder 2504 in a variety of ways. For example, the encoded enhancement layer video bitstream 530 and the encoded base layer video bitstream 532 may be provided over a bus to the video decoder 2504 or stored in memory for retrieval by the video decoder 2504.

- [0798] The video decoder 2504 may generate one or more decoded pictures based on the encoded enhancement layer video bitstream 530 and the encoded base layer video bitstream 532. A decoded picture 2118 (which may include an enhancement layer decoded picture 538 and a base layer decoded picture 540) may be displayed, played back, stored in memory and/or transmitted to another device, etc.
- [0799] In one example, the decoded picture 2118 may be transmitted to another device or back to the first electronic device 502a. The decoded picture 2118 may also be stored or otherwise maintained on the second electronic device 502b. In another example, the second electronic device 502b may display the decoded picture 2118. In other configurations, the decoded picture 2118 includes elements of the input picture 2506 with different properties based on the encoding and other operations performed on the bitstream 2114. In some configurations, the decoded picture 2118 may be included in a picture stream with a different resolution, format, specifications or other attribute from the input picture 2506.
- [0800] Figure 67B is a block diagram illustrating the use of a shared decoded picture buffer (DPB) 516c and a shared hybrid decoded picture buffer (DPB) operation module 520c for the base layer and the enhancement layer. Figure 67B includes the same components as that of Figure 67A except that the enhancement layer video decoder 534 and the base layer decoder 536 share both a decoded picture buffer (DPB) 516c and a hybrid decoded picture buffer (DPB) operation module 520c.
- [0801] Figure 68 is a timing diagram illustrating hybrid decoded picture buffer (DPB) 116 operation. The hybrid decoded picture buffer (DPB) 116 operations of Figure 68 show the steps of a preferred variant where removal, storage and reference marking are picture based and output, output marking and additional output are access unit (AU) based. An ideal decoded picture buffer (DPB) 116 and hypothetical reference decoder (HRD) may operate in such a manner that the various individual steps shown (e.g., removal 621a-b, output 623, storage 625a-c, marking 627a-c, 629 and additional output 631) are all performed instantaneously. The sequence of these steps and the timing offsets are illustrated between individual steps for illustration purposes. For output timing decoder conformance, the timing (relative to the delivery time of the first bit) of a picture output is the same for both the hypothetical reference decoder (HRD) and the decoder under test (DUT) up to a fixed delay. As such, the timing offset illustrated could happen for a DUT which introduces a fixed delay compared to an HRD.

- [0802] The steps for a decoded picture buffer (DPB) 116 of the first enhancement layer (EL1) 2642a, a decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b and a decoded picture buffer (DPB) 116 of the base layer (BL) 2644 are illustrated. A coded picture buffer (CPB) removal time 2646 is illustrated. After the coded picture buffer (CPB) removal time 2646, a picture based removal (without output) 621a may be performed by the decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b and a picture based removal (without output) 621b may be performed by the decoded picture buffer (DPB) 116 of the base layer (BL) 2644. After a timing offset, an access unit (AU) based picture output (bumping) 623 may be performed by the decoded picture buffer (DPB) 116 of the base layer (BL) 2644, the decoded picture buffer (DPB) 116 of the first enhancement layer (EL1) 2642a and the decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b.
- [0803] After another timing offset, the current decoded picture related process 613 is illustrated. A picture based storage step 625a-c may be performed by the decoded picture buffer (DPB) 116 of the base layer (BL) 2644, the decoded picture buffer (DPB) 116 of the first enhancement layer (EL1) 2642a and the decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b. The storage step may be performed after each of the base layer (BL) 2644, the first enhancement layer (EL1) 2642a and the second enhancement layer (EL2) 2642b pictures are decoded. The storage step 625 may be further subdivided. In the storage steps 625, a decoded picture is stored in the decoded picture buffer (DPB) 116 in an empty storage buffer and the decoded picture buffer (DPB) fullness is incremented by one. Also, when a picture is removed (without output) from the decoded picture buffer (DPB) 116, the decoded picture buffer (DPB) fullness is decremented by one. Similarly, when a picture is bumped from the decoded picture buffer (DPB) 116 (during either bumping or additional bumping), the decoded picture buffer (DPB) fullness is decremented by one.
- [0804] The decoded picture buffer (DPB) 116 may include separately identified and managed picture buffers for decoded pictures having different characteristics. For example, the decoded picture buffer (DPB) 116 may include separately identified and managed picture buffers for decoded pictures with different resolutions, different bit-depths and/or different color chromaticity.
- [0805] A decoded picture may instead be stored in a common pool of picture storage buffers in the decoded picture buffer (DPB) 116. For example, two additional sub-cases may be used to determine the decoded picture buffer (DPB) 116 size constraints that affect the bumping/removal process and level definitions. In a byte based decoded picture buffer (DPB) 116 constraint, a decoded picture may be stored with consideration for the size based on resolution and/or bit-depth. The decoded picture buffer (DPB) 116

size constraints may be defined as a byte limit that considers resolution and bit-depth of each decoded picture. In a picture unit based decoded picture buffer (DPB) 116 constraint, a decoded picture may be stored (and is considered to take one picture buffer slot). The decoded picture buffer (DPB) 116 size constraints may then be defined as a number of picture slots limit without considering resolution and bit-depth of each decoded picture.

- [0806] In one configuration, the decoded picture buffer (DPB) fullness may be tracked per layer. For example, the decoded picture buffer (DPB) 116 size constraints may be signaled, and bumping may be applied, per layer. Where each layer includes its own picture storage buffers a variable  $DPBFullness[nuh\_layer\_id]$  could be used to track to the decoded picture buffer (DPB) fullness of each layer. When a picture is removed from a layer with a layer ID value equal to  $nuh\_layer\_id$ , the variable  $DPBFullness[nuh\_layer\_id]$  may be set equal to  $DPBFullness[nuh\_layer\_id]-1$  (i.e.,  $DPBFullness[nuh\_layer\_id]$  may be decremented by one). In this case, the picture was removed from a picture storage buffer  $PSB[nuh\_layer\_id]$ .
- [0807] Similarly, when a currently decoded picture with a layer ID value equal to  $nuh\_layer\_id$  is stored in the decoded picture buffer (DPB) 116, the variable  $DPBFullness[nuh\_layer\_id]$  is set equal to  $DPBFullness[nuh\_layer\_id]+1$  (i.e.,  $DPBFullness[nuh\_layer\_id]$  is incremented by one). In this case, the picture was stored into a picture storage buffer  $PSB[nuh\_layer\_id]$ .
- [0808] The decoded picture buffer (DPB) fullness could also be tracked for an output layer set. The decoded picture buffer (DPB) 116 size constraints may then be signaled, and bumping may be applied, based on the constraints specified for an output layer set. A  $DPBFullness$  value could be tracked for the output layer set which is associated with the operation point under test. Thus, when a picture is removed from a layer belonging to the output layer set, the value of the decoded picture buffer (DPB) fullness may be decremented by one as  $DPBFullness = DPBFullness-1$ . Likewise, when a currently decoded picture is stored in the decoded picture buffer (DPB) 116, the decoded picture buffer (DPB) fullness may be decremented by one as  $DPBFullness = DPBFullness+1$ .
- [0809] Within the current decoded picture related process 613, the marking process 611 is illustrated. The marking process 611 may include reference marking 627a-c and output marking 629. After a timing offset from the storage steps 625a-c, a picture based reference marking step 627a-c may be performed by the decoded picture buffer (DPB) 116 of the base layer (BL) 2644, the decoded picture buffer (DPB) 116 of the first enhancement layer (EL1) 2642a and the decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b. After another timing offset, an access unit (AU) based output marking step 629 may be performed by the decoded picture buffer (DPB) 116 of the base layer (BL) 2644, the decoded picture buffer (DPB) 116 of the first en-

hancement layer (EL1) 2642a and the decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b. Once the current decoded picture related process 613 is completed, an access unit (AU) based picture output (additional bumping) step 631 may be performed by the decoded picture buffer (DPB) 116 of the base layer (BL) 2644, the decoded picture buffer (DPB) 116 of the first enhancement layer (EL1) 2642a and the decoded picture buffer (DPB) 116 of the second enhancement layer (EL2) 2642b.

- [0810] Figure 69 is a block diagram illustrating structure and timing for network abstraction layer (NAL) units of layers for coded pictures and access units (AUs) when the second enhancement layer (EL2) 942b has a lower picture rate than the base layer (BL) 944 and the first enhancement layer (EL1) 942a. The NAL units of an EL1 coded picture 953a are illustrated along the first enhancement layer (EL1) 942a. The NAL units of an EL2 coded picture 953b are illustrated along the second enhancement layer (EL2) 942b. The NAL units of a base layer coded picture 953c are illustrated along the base layer (BL) 944.
- [0811] At time t1, the NAL units of an EL1 coded picture 953a, the NAL units of an EL2 coded picture 953b and the NAL units of a base layer coded picture 953c are part of the access unit (AU) 955a. At time t2, the NAL units of an EL1 coded picture 953a and the NAL units of a base layer coded picture 953c are part of the access unit (AU) 955b. At time t3, the NAL units of an EL1 coded picture 953a, the NAL units of an EL2 coded picture 953b and the NAL units of a base layer coded picture 953c are part of the access unit (AU) 955c. At time t4, the NAL units of an EL1 coded picture 953a and the NAL units of a base layer coded picture 953c are part of the access unit (AU) 955d.
- [0812] Figure 70 is a block diagram illustrating structure and timing for network abstraction layer (NAL) units of layers for coded pictures and access units (AUs) when the base layer (BL) 1044 has a lower picture rate than the first enhancement layer (EL1) 1042a and the second enhancement layer (EL2) 1042b. The NAL units of an EL1 coded picture 1053a are illustrated along the first enhancement layer (EL1) 1042a. The NAL units of an EL2 coded picture 1053b are illustrated along the second enhancement layer (EL2) 1042b. The NAL units of a base layer coded picture 1053c are illustrated along the base layer (BL) 1044.
- [0813] At time t1, the NAL units of an EL1 coded picture 1053a, the NAL units of an EL2 coded picture 1053b and the NAL units of a base layer coded picture 1053c are part of the access unit (AU) 1055a. At time t2, the NAL units of an EL1 coded picture 1053a and the NAL units of a EL2 coded picture 1053b are part of the access unit (AU) 1055b. At time t3, the NAL units of an EL1 coded picture 1053a, the NAL units of an EL2 coded picture 1053b and the NAL units of a base layer coded picture 1053c are part of the access unit (AU) 1055c. At time t4, the NAL units of an EL1 coded picture



1053a and the NAL units of an EL1 coded picture 1053b are part of the access unit (AU) 1055d.

- [0814] The term "computer-readable medium" refers to any available medium that can be accessed by a computer or a processor. The term "computer-readable medium," as used herein, may denote a computer- and/or processor-readable medium that is non-transitory and tangible. By way of example, and not limitation, a computer-readable or processor-readable medium may comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium that can be used to carry or store desired program code in the form of instructions or data structures and that can be accessed by a computer or processor. Disk and disc, as used herein, includes compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk and Blu-ray (registered trademark) disc where disks usually reproduce data magnetically, while discs reproduce data optically with lasers.
- [0815] It should be noted that one or more of the methods described herein may be implemented in and/or performed using hardware. For example, one or more of the methods or approaches described herein may be implemented in and/or realized using a chipset, an ASIC, a LSI or integrated circuit, etc.
- [0816] Each of the methods disclosed herein comprises one or more steps or actions for achieving the described method. The method steps and/or actions may be interchanged with one another and/or combined into a single step without departing from the scope of the claims. In other words, unless a specific order of steps or actions is required for proper operation of the method that is being described, the order and/or use of specific steps and/or actions may be modified without departing from the scope of the claims.
- [0817] It is to be understood that the claims are not limited to the precise configuration and components illustrated above. Various modifications, changes and variations may be made in the arrangement, operation and details of the systems, methods, and apparatus described herein without departing from the scope of the claims.

CLAIMS:

1. A method for video decoding, comprising:
  - receiving a bitstream comprising a video parameter set and bits representing a current picture;
  - parsing at least a portion of a slice header of the current picture;
  - performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture;
  - performing a second determination, separate from the first determination, based on one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis;
  - performing the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis;
  - performing the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis;
  - performing a decoding and storing of a current decoded picture in the DPB;
  - performing a third determination, based on the one or more syntax elements in the video parameter set, whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis;

marking the current decoded picture in the DPB in accordance with the third determination whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis; and

performing an additional picture output from the DPB in accordance with a fourth determination whether the additional picture output is to be performed on a per-picture basis or per-AU basis.

2. The method of claim 1, wherein the removal and the output from the DPB is based on at least one AU output flag.
3. The method of claim 2, wherein the AU output flag is one of an AU output flag, an AU no Random Access Skipped Leading (RASL) output flag and an AU no output of prior pictures flag.
4. The method of claim 2 or 3, wherein the AU output flag is obtained from a slice header.
5. The method of any one of claims 1 to 4, wherein all picture output flag syntax element values are constrained to the same value for all coded pictures in an AU when a picture output flag is present in the slice header.
6. The method of any one of claims 1 to 5, wherein the removal is performed on a per-picture basis, wherein the picture output is performed on a per-AU basis, wherein the storing and decoding are on a per-picture basis, wherein the marking is on a per-picture basis, and wherein the additional picture output is performed on a per-AU basis.
7. The method of any one of claims 1 to 5, wherein the removal, the picture output, the decoding and storing, the marking and the additional picture output are performed on a per-AU basis.

8. The method of any one of claims 1 to 5, wherein the removal, the decoding and storing and the marking are performed on a per-picture basis, and wherein the picture output and the additional picture output are performed on a per-AU basis.

9. The method of any one of claims 1 to 5, wherein the removal, the picture output, the decoding and storing, the marking and the additional picture output are performed on a per-picture basis.

10. The method of any one of claims 1 to 9, wherein the DPB comprises separately identified and managed picture buffers for decoded pictures having one or more of different resolutions, different bit-depths and different color chromaticity.

11. An electronic device configured for video coding, comprising:

a processor; and

memory in electronic communication with the processor, wherein instructions stored in the memory are executable to:

receive a bitstream comprising a video parameter set and bits representing a current picture,

parse at least a portion of a slice header of the current picture,

performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture,

perform a second determination, separate from the first determination, based on the one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis,

perform the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis,

perform the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis,

perform a decoding and storing of a current decoded picture in the DPB,

perform a third determination, based on the one or more syntax elements in the video parameter set, whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis,

mark the current decoded picture in the DPB in accordance with the third determination whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis, and

perform an additional picture output from the DPB in accordance with a fourth determination whether the additional picture output is to be performed on a per-picture basis or per-AU basis.

12. The electronic device of claim 11, wherein the removal and the output from the DPB is based on at least one AU output flag.

13. The electronic device of claim 12, wherein the AU output flag is one of an AU output flag, an AU no Random Access Skipped Leading (RASL) output flag and an AU no output of prior pictures flag.

14. The electronic device of claim 12 or 13, wherein the AU output flag is obtained from a slice header.

15. The electronic device of any one of claims 11 to 14, wherein all picture output flag syntax element values are constrained to the same value for all coded pictures in an AU when a picture output flag is present in the slice header.

16. The electronic device of any one of claims 11 to 15, wherein the removal is performed on a per-picture basis, wherein the picture output is performed on a per-AU basis, wherein the storing and decoding are on a per-picture basis, wherein the marking is on a per-picture basis, and wherein the additional picture output is performed on a per-AU basis.

17. The electronic device of any one of claims 11 to 15, wherein the removal, the picture output, the decoding and storing, the marking and the additional picture output are performed on a per-AU basis.

18. The electronic device of any one of claims 11 to 15, wherein the removal, the decoding and storing and the marking are performed on a per-picture basis, and wherein the picture output and the additional picture output are performed on a per-AU basis.

19. The electronic device of any one of claims 11 to 15, wherein the removal, the picture output, the decoding and storing, the marking and the additional picture output are performed on a per-picture basis.

20. The electronic device of any one of claims 11 to 19, wherein the DPB comprises separately identified and managed picture buffers for decoded pictures having one or more of different resolutions, different bit-depths and different color chromaticity.

21. A method for video decoding, comprising:

receiving a bitstream comprising a video parameter set and bits representing a current picture;

parsing at least a portion of a slice header of the current picture;

performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to

be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture;

performing a second determination, separate from the first determination, based on one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis;

performing the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis;

performing the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis;

performing a decoding and storing of a current decoded picture in the DPB.

22. The method of claim 21, wherein the removal and the output from the DPB is based on at least one AU output flag.

23. The method of claim 22, wherein the AU output flag is one of an AU output flag, an AU no Random Access Skipped Leading (RASL) output flag and an AU no output of prior pictures flag.

24. The method of claim 22 or 23, wherein the AU output flag is obtained from a slice header.

25. The method of claim 24, wherein all picture output flag syntax element values are constrained to the same value for all coded pictures in an AU when a picture output flag is present in the slice header.

26. The method of any one of claims 21 to 25, wherein the removal is performed on a per-picture basis, wherein the picture output is performed on a per-AU basis, wherein the storing and decoding are on a per-picture basis.

27. The method of any one of claims 21 to 25, wherein the removal, the picture output, and the decoding and storing, are performed on a per-AU basis.
28. The method of any one of claims 21 to 25, wherein the removal, the decoding and storing are performed on a per-picture basis, and wherein the picture output is performed on a per-AU basis.
29. The method of any one of claims 21 to 25, wherein the removal, the picture output, and the decoding and storing are performed on a per-picture basis.
30. The method of any one of claims 21 to 29, wherein the DPB comprises separately identified and managed picture buffers for decoded pictures having one or more of different resolutions, different bit-depths and different color chromaticity.
31. The method of any one of claims 21 to 30, further comprising:
- performing a third determination, based on the one or more syntax elements in the video parameter set, whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis; and
- marking the current decoded picture in the DPB in accordance with the third determination whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis.
32. The method of any one of claims 21 to 30, further comprising:
- performing an additional picture output from the DPB in accordance with a third determination whether the additional picture output is to be performed on a per-picture basis or per-AU basis.
33. An electronic device configured for video coding, comprising:
- a processor; and



memory in electronic communication with the processor, wherein instructions stored in the memory are executable to:

receive a bitstream comprising a video parameter set and bits representing a current picture,

parse at least a portion of a slice header of the current picture,

performing a first determination, based on one or more syntax elements in the video parameter set, whether a removal operation from a decoded picture buffer (DPB) is to be performed on a per-picture basis or per-access unit (AU) basis, wherein an access unit comprises at least a base layer and an enhanced layer, and each layer represents a picture,

perform a second determination, separate from the first determination, based on the one or more syntax elements in the video parameter set, whether a picture output operation from the DPB is to be performed on a per-picture basis or per-AU basis,

perform the removal operation from the DPB in accordance with the first determination whether the removal operation is to be performed on a per-picture basis or per-AU basis,

perform the picture output operation from the DPB in accordance with the second determination whether the picture output operation is to be performed on a per-picture basis or per-AU basis, and

perform a decoding and storing of a current decoded picture in the DPB.

34. The electronic device of claim 33, wherein the removal and the output from the DPB is based on at least one AU output flag.

35. The electronic device of claim 34, wherein the AU output flag is one of an AU output flag, an AU no Random Access Skipped Leading (RASL) output flag and an AU no output of prior pictures flag.

36. The electronic device of claim 34 or 35, wherein the AU output flag is obtained from a slice header.

37. The electronic device of claim 36, wherein all picture output flag syntax element values are constrained to the same value for all coded pictures in an AU when a picture output flag is present in the slice header.

38. The electronic device of any one of claims 33 to 37, wherein the removal is performed on a per-picture basis, wherein the picture output is performed on a per-AU basis, and wherein the storing and decoding are on a per-picture basis.

39. The electronic device of any one of claims 33 to 37, wherein the removal, the picture output, and the decoding and storing are performed on a per-AU basis.

40. The electronic device of any one of claims 33 to 37, wherein the removal, and the decoding and storing are performed on a per-picture basis, and wherein the picture output is performed on a per-AU basis.

41. The electronic device of any one of claims 33 to 37, wherein the removal, the picture output, and the decoding and storing are performed on a per-picture basis.

42. The electronic device of any one of claims 33 to 41, wherein the DPB comprises

separately identified and managed picture buffers for decoded pictures having one or more of different resolutions, different bit-depths and different color chromaticity.

43. The electronic device of any one of claims 33 to 42, wherein the instructions stored in the memory are executable to:

perform a third determination, based on the one or more syntax elements in the video parameter set, whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis; and

mark the current decoded picture in the DPB in accordance with the third determination whether marking the current decoded picture is to be performed on a per-picture basis or per-AU basis.

44. The electronic device of any one claims 33 to 42, wherein the instructions stored in the memory are executable to:

perform an additional picture output from the DPB in accordance with a third determination whether the additional picture output is to be performed on a per-picture basis or per-AU basis.

[Fig. 1A]

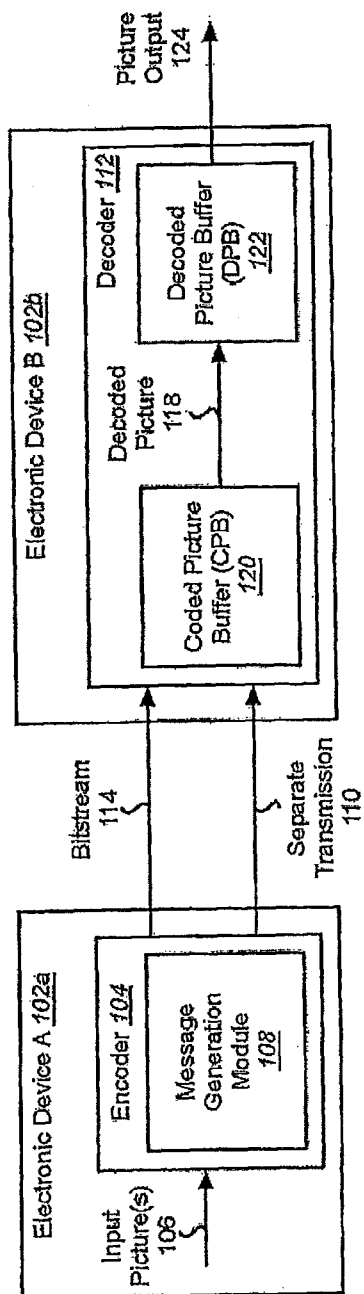


FIG. 1A

[Fig. 1B]

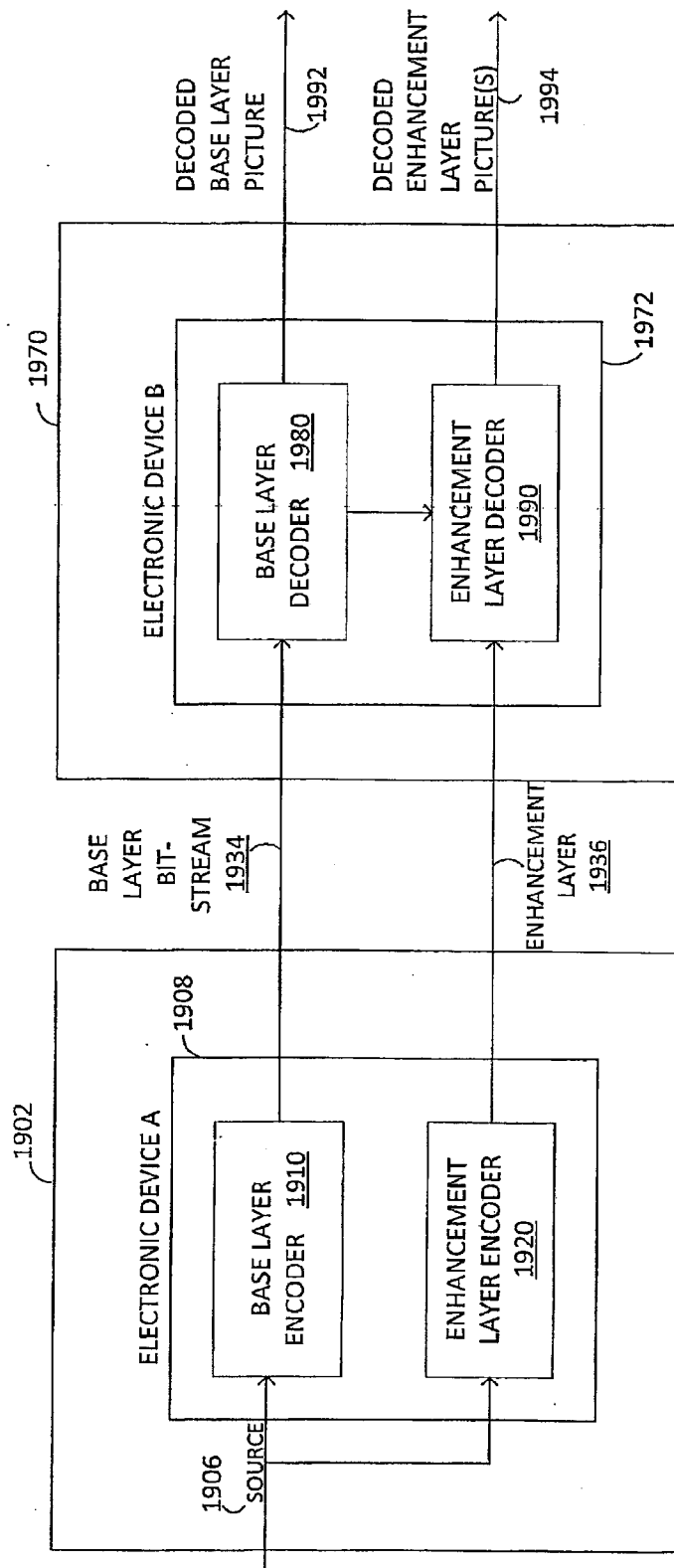


FIG. 1B

[Fig. 2]

200

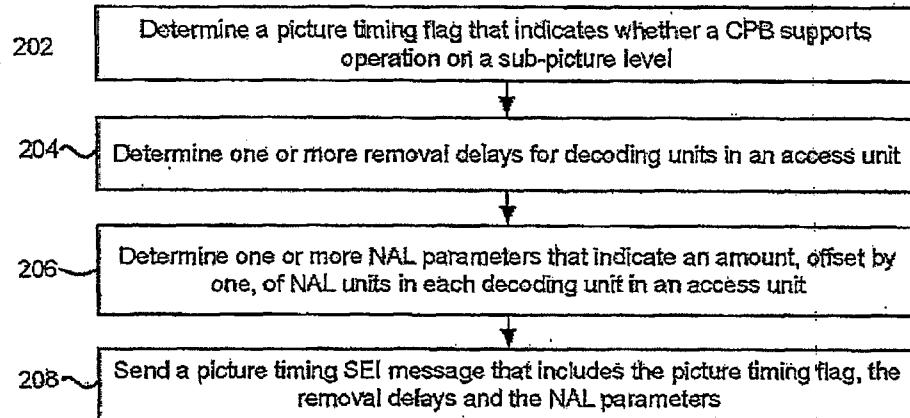


FIG. 2

[Fig. 3]

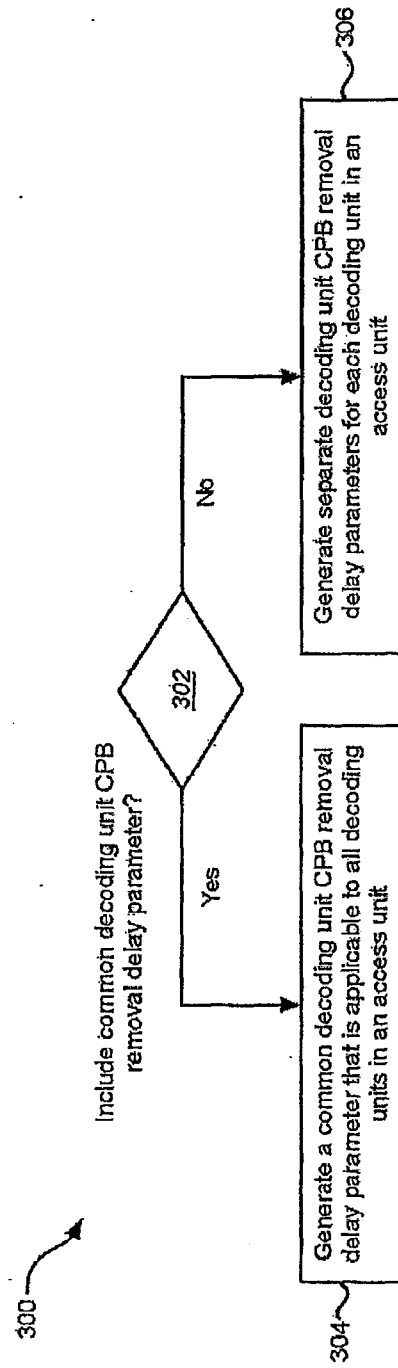


FIG. 3

[Fig. 4]

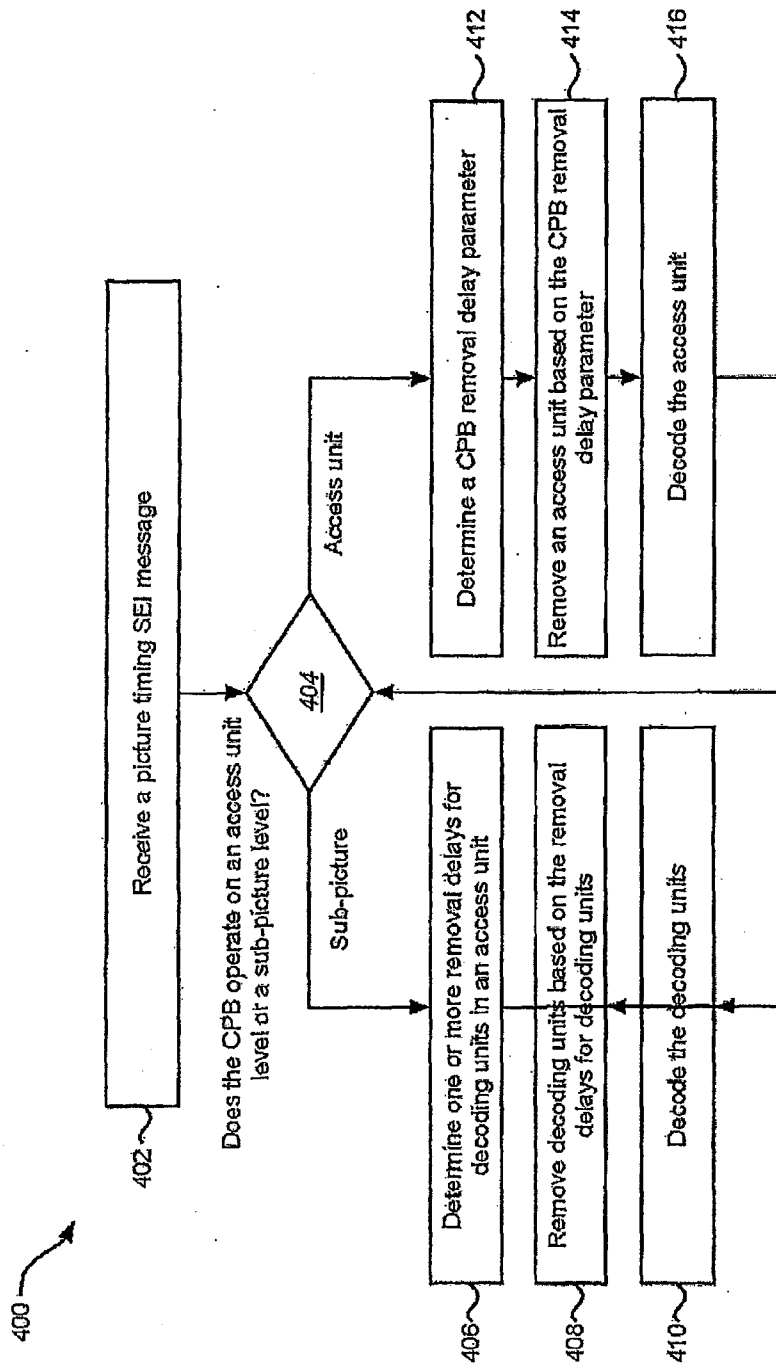


FIG. 4



[Fig. 5]

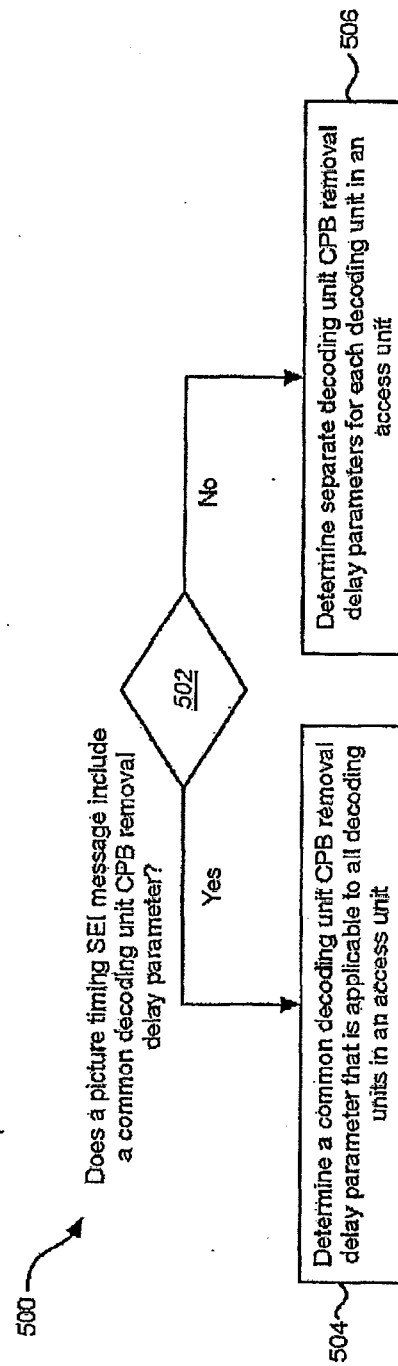
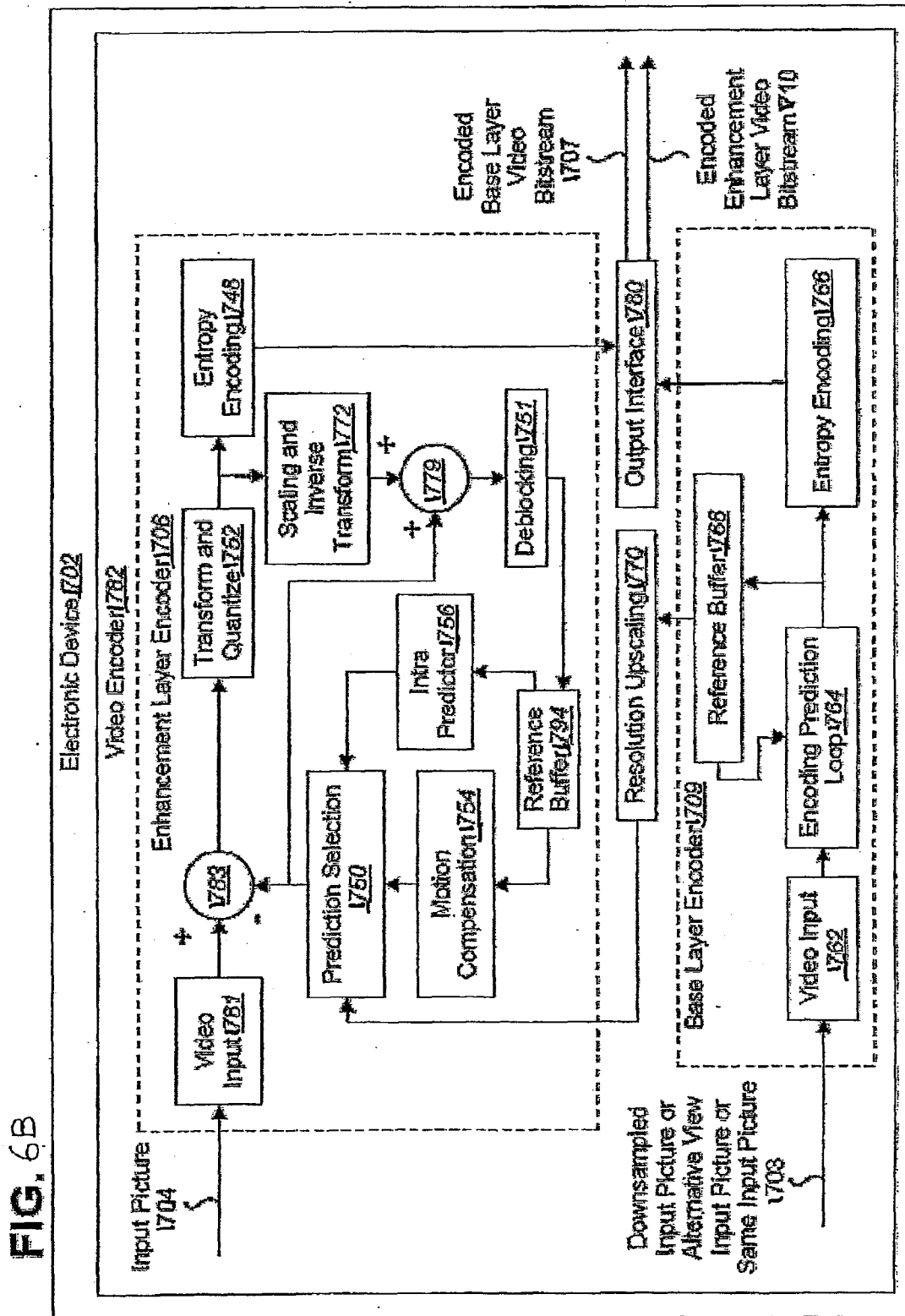


FIG. 5

The diagram illustrates the architecture of an encoder 604 and a decoder 606. The encoder 604 processes input picture(s) 606 from source 622. It includes an intra-frame prediction module and reconstruction buffer 624, a motion estimation and motion compensation module 646, a transformation module 632, a quantization module 636, an entropy coding module 642, and a message generation module 608. The encoder produces two bitstreams: Bitstream A 614a and Bitstream B 614b. The decoder 606 reconstructs the pictures from these bitstreams. It includes an ALF module 672, a decoded picture buffer 676, an SAO module 666, a deblocking filter 662, a reconstruction module 658, and an inverse transformation module 654. The decoder also includes an inverse quantization module 650 and an inverse entropy coding module 642. The decoder produces a reference picture 678 and a decoded picture 676.

**FIG. 6A**

[Fig. 6B]



[Fig. 7A]

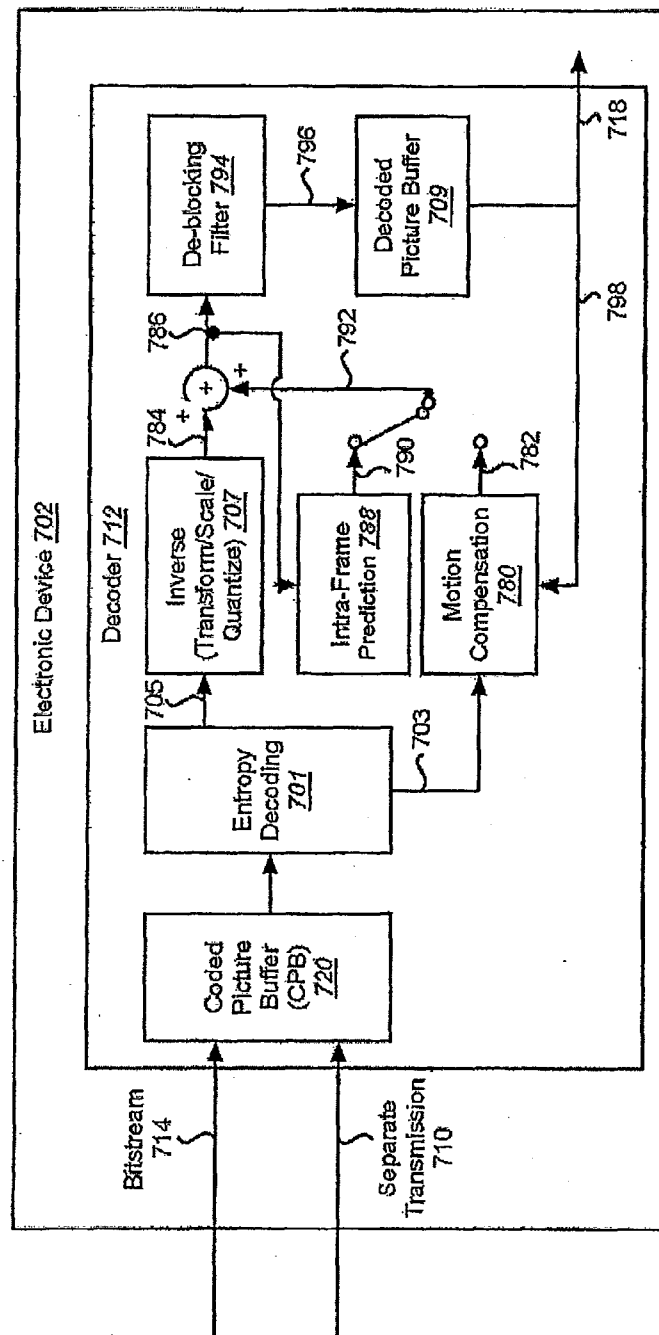
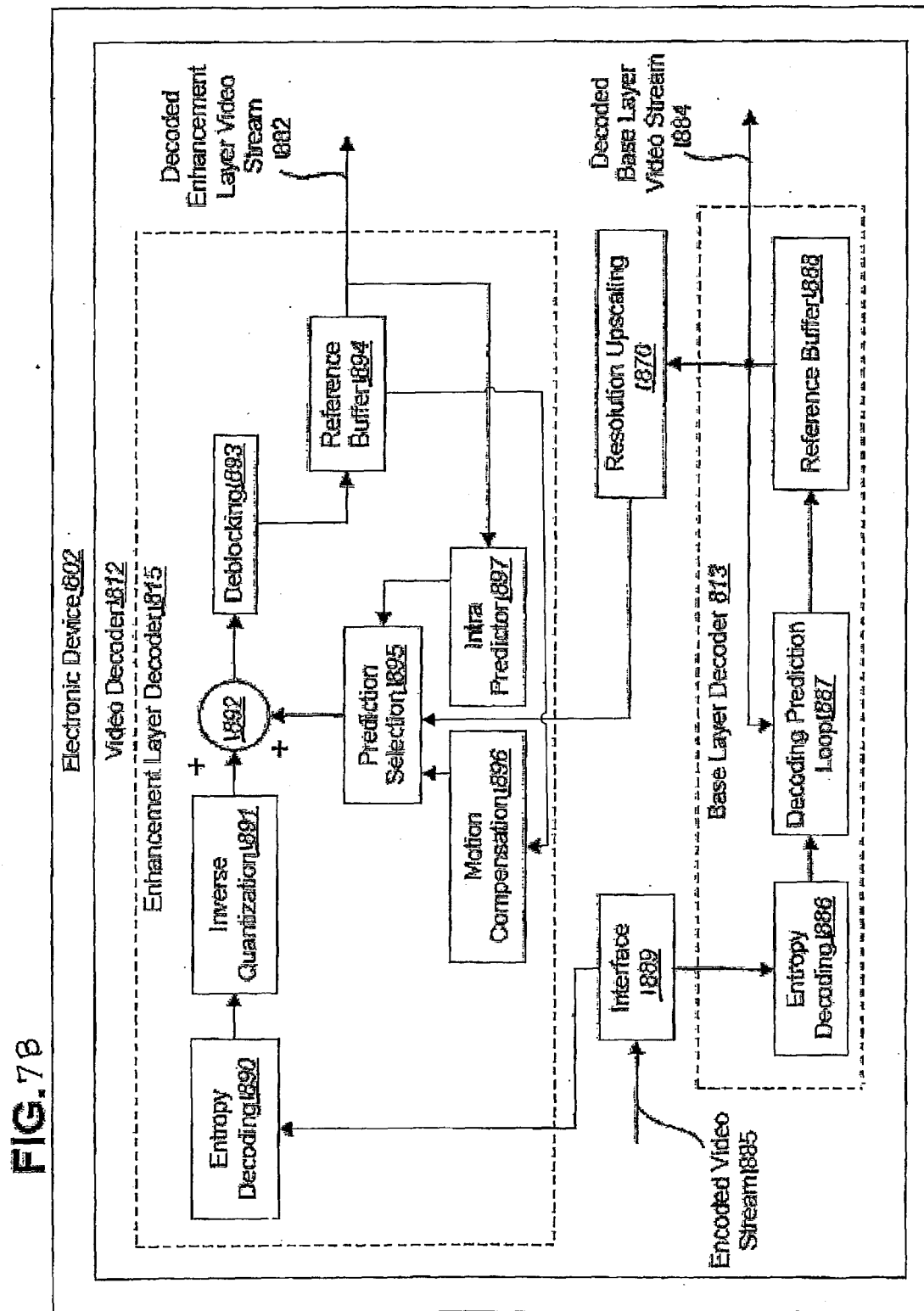


FIG. 7A

[Fig. 7B]



[Fig. 8]

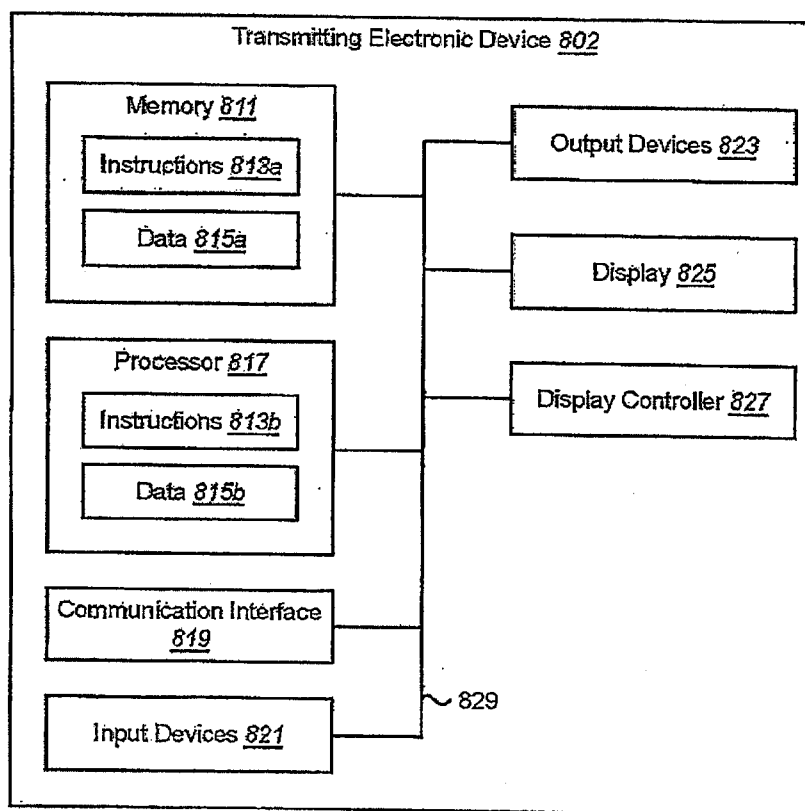


FIG. 8

[Fig. 9]

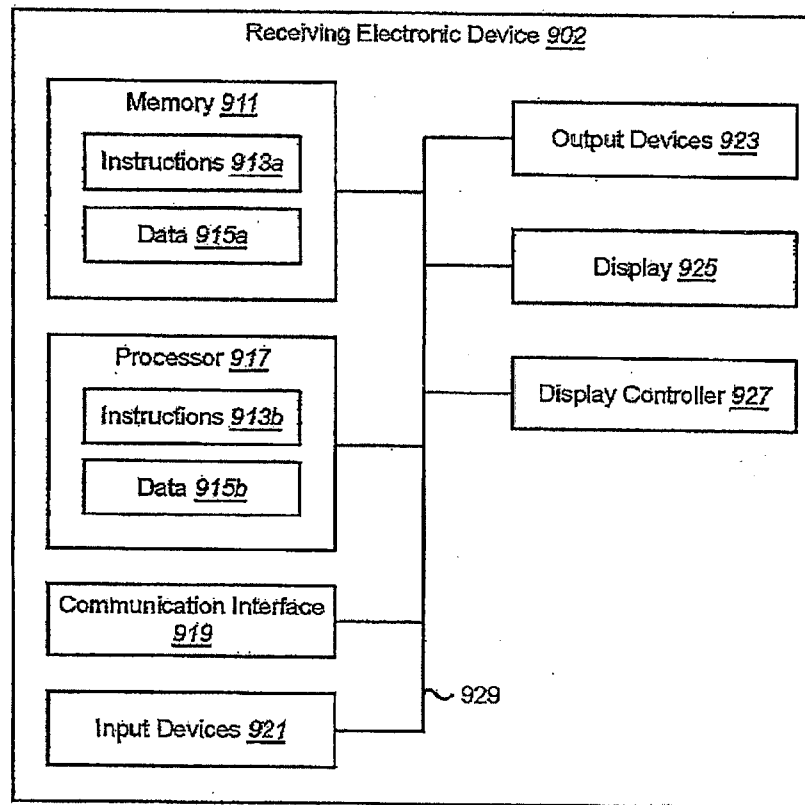


FIG. 9

[Fig. 10]

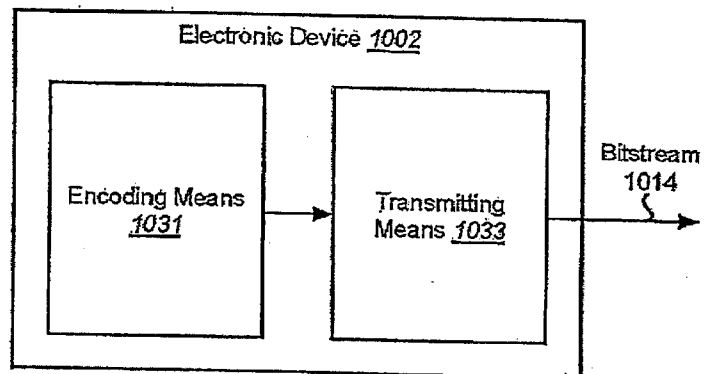


FIG. 10

[Fig. 11]

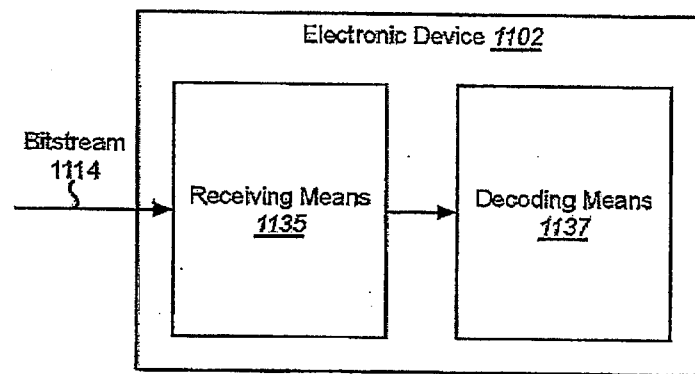


FIG. 11

[Fig. 12]

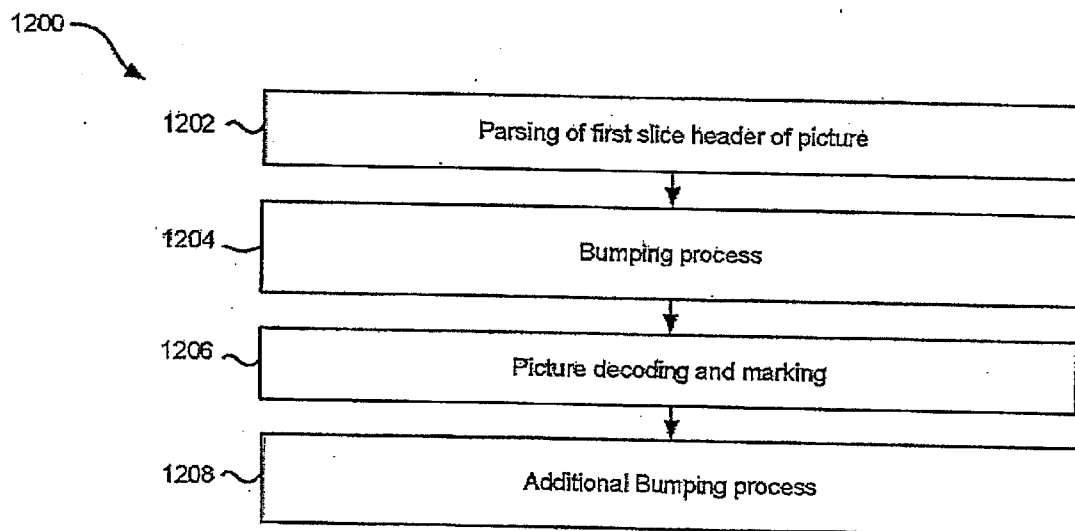


FIG. 12

[Fig. 13A]

nal_unit_header() {	
forbidden_zero_bit	f(1)
nal_unit_type	u(6)
nuh_reserved_zero_6bits	u(6)
nuh_temporal_id_plus1	u(3)
}	

NAL UNIT HEADER SYNTAX

FIG. 13A



[Fig. 13B]

nal_unit_header() {	Descriptor
forbidden_zero_bit	f(1)
nal_unit_type	u(6)
layer_id_plus1	u(6)
nuh_temporal_id_plus1	u(3)
}	

NAL UNIT HEADER SYNTAX

FIG. 13B

[Fig. 13C]

nal_unit_header() {	Descriptor
forbidden_zero_bit	f(1)
nal_unit_type	u(6)
layer_id	u(6)
nuh_temporal_id_plus1	u(3)
}	

NAL UNIT HEADER SYNTAX

FIG. 13C

[Fig. 14]

nal_unit(NumBytesInNALunit) {	Descriptor
nal_unit_header() .	
NumBytesInRBSP == 0	
for( i = 2; i < NumBytesInNALunit; i++ ) {	
if( i + 2 < NumBytesInNALunit && next_bits( 24 ) == 0x000003 ) {	
rbsp_byte[ NumBytesInRBSP++ ]	b(8)
rbsp_byte[ NumBytesInRBSP++ ]	b(8)
i += 2	
emulation_prevention_three_byte /* equal to 0x03 */	f(8)
} else	
rbsp_byte[ NumBytesInRBSP++ ]	b(8)
}	
}	

GENERAL NAL UNIT SYNTAX

FIG. 14

[Fig. 15]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_zero_bit	u(1)
// scalability type and layer_id partitioning method	
scalability_type	u(4)
for( i=0; i < MaxDim( scalability_type ); i++ )	
layer_id_dim_len[ i ]	u(3)
// layer specific information	
for( i=0; i <= max_num_layers_minus1; i++ ) {	
vps_layer_id[ i ]	u(6)
// layer dependency	
num_direct_ref_layers[ i ]	u(6)
for( j=0; j < num_direct_ref_layers[ i ]; j++ )	
ref_layer_id[ i ][ j ]	u(6)
}	
}	

Existing video parameter set extension syntax

FIG. 15

[Fig. 16]

scalability type	MaxDim(scalability_type)	Scalability dimensions
0	1	none (base H.264)
1	2	spatial and quality
2	3	spatial, quality, unspecified
3	4	spatial, quality, unspecified, unspecified
4	2	multiview and depth
5	3	multiview, depth, unspecified
6	4	multiview, depth, unspecified, unspecified
7	4	multiview, spatial, quality and depth
8	5	multiview, spatial, quality, depth, unspecified
9	6	multiview, spatial, quality, depth, unspecified, unspecified
10...15	reserved	reserved

Existing Scalability Types

FIG. 16

[Fig. 17]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_zero_bit	u(1)
// Scalability map and layer_id_plus1 mapping method	
for (i=0; i<=max_num_layers_minus1; i++) {	
scalability_map[i]	u(3)
}	
// layer specific information	
for( i = 0; i <= max_num_layers_minus1; i++) {	
vps_layer_id[ i ]	u(6)
// layer dependency	
num_direct_ref_layers[ i ]	u(6)
for( j = 0; j < num_direct_ref_layers[ i ]; j++)	
ref_layer_id[ i ][ j ]	u(6)
}	
}	

Video parameter set extension syntax

FIG. 17

[Fig. 18]

Scalability_map[i]	Scalability dimension
0	nons (base HEVC)
1	spatial
2	quality
3	depth
4	multiview
5	unspecified
6, 7	reserved

Scalability Map Syntax

FIG. 18

[Fig. 19]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_zero_bit	u(1)
// scalability map and layer_id_plus1 mapping method	
num_scalability_dimensions_minus1	u(3)
for( i=0; i<= num_scalability_dimensions_minus1; i++) {	
scalability_map[i]	u(3)
num_bits_for_scalability_map[i]	u(3)
}	
// layer specific information	
for( i=0; i<= max_num_layers_minus1; i++) {	
vps_layer_id[i]	u(6)
// layer dependency	
num_direct_ref_layers[i]	u(6)
for( j=0; j< num_direct_ref_layers[i]; j++)	
ref_layer_id[i][j]	u(6)
}	
}	

Video parameter set extension syntax

FIG. 19

[Fig. 20]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_zero_bit	u(1)
// layer specific information	
for( i=1; i<= vps_max_layers_minus1; i++) {	
// mapping of layer ID to scalability dimension IDs	
num_dimensions_minus1[i]	u(4)
for( j=0; j<= num_dimensions_minus1; j++) {	
dimension_type[i][j]	u(4)
dimension_id[i][j]	u(8)
}	
// layer dependency	
num_direct_ref_layers[i]	u(6)
for( j=0; j< num_direct_ref_layers[i]; j++)	
ref_layer_id[i][j]	u(6)
}	
}	

Existing video parameter set extension syntax

FIG. 20

[Fig. 21]

dimension_type[i][j]	dimension_id[i][j]
0	view order idx
1	depth flag
2	dependency ID
3	quality ID
4..15	reserved

Existing Dimension Type, Dimension ID Syntax

FIG. 21

[Fig. 22]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_zero_bit	u(1)
// layer specific information	
for( i = 1; i <= max_num_layers_minus1; i++ ) {	
// mask signalling scalability types that are present for this layer ID	
scalability_mask	u(8)
for( j = 0; j <= num_scalability_types[i]; j++ ) {	
scalability_id[j]	u(8)
}	
// layer dependency	
num_direct_ref_layers[i]	u(6)
for( j = 0; j < num_direct_ref_layers[i]; j++ )	
ref_layer_id[i][j]	u(6)
}	
}	

Video parameter set extension syntax

FIG. 22

[Fig. 23]

scalability_mask_bit[k]	Scalability dimension
0	spatial
1	quality
2	depth
3	multiview
4-15	Reserved

Scalability Map Syntax

FIG. 23

[Fig. 24]

video_parameter_set_rbsp() {	Descriptor
video_parameter_set_id	u(4)
vps_temporal_id_nesting_flag	u(1)
reserved_zero_2bits	u(2)
max_num_layers_minus1 // reserved_zero_6bits	u(6)
vps_max_sub_layers_minus1	u(3)
profile_level(1, vps_max_sub_layers_minus1)	
next_essential_info_byte_offset // reserved_zero_12bits	u(12)
for(i = 0; i <= vps_max_sub_layers_minus1; i++) {	
vps_max_dec_pic_buffering[i]	ue(v)
vps_max_num_reorder_pics[i]	ue(v)
vps_max_latency_increase[i]	ue(v)
}	
num_hrd_parameters	ue(v)
for(i = 0; i < num_hrd_parameters; i++) {	
if(i > 0)	
op_point(i)	
hrd_parameters(i == 0, vps_max_sub_layers_minus1)	
}	
bit_equal_to_one	u(1)
scalability_mask	u(8)
for(i = 0; i <= max_num_layers_minus1; i++) {	
vps_extension_data()	
}	
vps_extension_flag	u(1)
if(vps_extension_flag)	
while(more_rbsp_data())	
vps_extension_data_flag	u(1)
rbsp_trailing_bits()	
}	
}	

Video parameter set extension syntax

FIG. 24

[Fig. 25]

vps_extension_data() {	Descriptor
vps_extension_type	u(2)
layer_id	u(6)
layer_max_sub_layers_minus1	u(3)
new_profile_level_flag	u(1)
if (new_profile_level_flag == 0) {	
profile_reference_layer_id	u(6)
}	
else {	
profile_level(1, layer_max_sub_layers_minus1)	
}	
for (j = 0; j <= num_scalability_types; j++) {	
scalable_id[j]	u(8)
}	
num_reference_layers	u(6)
for (n = 0; n < num_reference_layers; n++)	
direct_coding_dependency_layer_id_plus1[n]	u(6)
}	

Video parameter set extension syntax

FIG. 25

[Fig. 26]

vps_extension() {	Descriptor
while( byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_for_minus1[i]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability-dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[i]	u(6)
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[i][j]	u(v8)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_for_level( 1, vps_max_sub_layers_minus1 )	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// layer dependency	
num_direct_ref_layers[i]	u(4)
for( j = 0; j < num_direct_ref_layers[i]; j++ )	
ref_layer_id[i][j]	u(6)
}	
}	

Video Parameter Set Extension Syntax

FIG. 26

[Fig. 27]

scalability_mask's k'th bit	Scalability dimension
0	spatial
1	quality
2	depth
3	multiview
4-15	Reserved

Scalability\_mask Syntax

FIG. 27



[Fig. 28]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[ i ]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_tier_level( i, vps_max_sub_layers_minus1 )	
layer_dependency_information_pattern	u(v)
for( i = 1; i <= NumDepLayers; i++ ) {	
// layer dependency information signaling	
num_direct_ref_layers[ i ]	u(6)
for( j = 0; j < num_direct_ref_layers[ i ]; j++ )	
ref_layer_id[ i ][ j ]	u(6)
}	
}	

Video Parameter Set Extension Syntax

FIG. 28

[Fig. 29]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[i]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[i]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[i][j]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_tier_level( i, vps_max_sub_layers_minus1 )	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// layer dependency information signaling	
layer_dependency_flag[i]	u(1)
if( layer_dependency_flag[i] ) {	
num_direct_ref_layers[i]	u(5)
for( j = 0; j < num_direct_ref_layers[i]; j++ )	
ref_layer_id[i][j]	u(6)
}	
}	
}	

Video Parameter Set Extension Syntax

FIG. 29

[Fig. 30]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[ i ]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_tier_level( i, vps_max_sub_layers_minus1 )	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// layer dependency information signaling	
layer_dependency_map[ i ]	u(v)
}	
}	

Video Parameter Set Extension Syntax

FIG. 30

[Fig. 31]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
ave_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[i]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_tier_level( 1, vps_max_sub_layers_minus1 )	
layer_dependency_information_pattern	u(v)
for( i = 1; i <= NumDepLayers; i++ )	
layer_dependency_map[i]	u(v)
}	

Video Parameter Set Extension Syntax

FIG. 31

[Fig. 32]

vps_extension() {	Descriptor
while( lbyte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[ i ]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_tier_level( i, vps_max_sub_layers_minus1 )	
layer_dependency_information_pattern	u(v)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( layer_dependency_information_pattern( i ) ) {	
// layer dependency information signaling	
num_direct_ref_layers[ i ]	u(6)
for( j = 0; j < num_direct_ref_layers[ i ]; j++ )	
ref_layer_id[ i ][ j ]	u(6)
}	
}	
}	

Video Parameter Set Extension Syntax

FIG. 32.

[Fig. 33]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_codec_flag	u(1)
scalability_mask	u(16)
for( i = 0; i < NumScalabilityTypes; i++ ) {	
dimension_id_len_minus1[ i ]	u(3)
}	
vps_nuh_layer_id_present_flag	u(1)
// layer specific information	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
// mapping of layer ID to scalability dimension IDs	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j <= num_dimensions_minus1; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
profile_tier_level( 1, vps_max_sub_layers_minus1 )	
layer_dependency_information_pattern	u(v)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( layer_dependency_information_pattern( i ) ) {	
layer_dependency_map[ i ]	u(v)
}	
}	

Video Parameter Set Extension Syntax

FIG. 33

[Fig. 34]

video_parameter_set_rbsp() {	Descriptor
vps_video_parameter_set_id	u(4)
vps_reserved_three_2bits	u(2)
vps_max_layers_minus1	u(6)
vps_max_sub_layers_minus1	u(3)
vps_temporal_id_nesting_flag	u(1)
vps_extension_offset // vps_reserved 0xffff_16bits	u(16)
profile_tier_level( 1, vps_max_sub_layers_minus1 )	
vps_sub_layer_ordering_info_present_flag	u(1)
for( i = ( vps_sub_layer_ordering_info_present_flag ? 0 : vps_max_sub_layers_minus1 ); i <= vps_max_sub_layers_minus1; i++ ) {	
vps_max_dec_pic_buffering_minus1[ i ]	ue(v)
vps_max_num_reorder_pics[ i ]	ue(v)
vps_max_latency_increase_plus1[ i ]	ue(v)
}	
vps_max_layer_id	u(6)
vps_num_layer_sets_minus1	ue(v)
for( i = 1; i <= vps_num_layer_sets_minus1; i++ )	
for( j = 0; j <= vps_max_layer_id; j++ )	
layer_id_included_flag[ i ][ j ]	u(1)
vps_timing_info_present_flag	u(1)
if( vps_timing_info_present_flag ) {	
vps_num_units_in_tick	u(32)
vps_time_scale	u(32)
vps_poc_proportional_to_timing_flag	u(1)
if( vps_poc_proportional_to_timing_flag )	
vps_num_ticks_poc_diff_one_minus1	ue(v)
vps_num_hrd_parameters	ue(v)
for( i = 0; i < vps_num_hrd_parameters; i++ ) {	
hrd_layer_set_idx[ i ]	ue(v)
if( i > 0 )	
cprms_present_flag[ i ]	u(1)
hrd_parameters( cprms_present_flag[ i ], vps_max_sub_layers_minus1 )	
}	
}	
vps_extension_flag	u(1)
if( vps_extension_flag ) {	
vps_extension()	
vps_extension2_flag	u(1)
if( vps_extension2_flag )	
while( more_rbsp_data() )	
vps_extension_data_flag	u(1)
}	
rbsp_trailing_bits()	
}	

FIG. 34

[Fig. 35]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_layer_flag	u(1)
splitting_flag	u(1)
for( i = 0; NumScalabilityTypes = 0; i < 16; i++ ) {	
scalability_mask[ i ]	u(1)
NumScalabilityTypes += scalability_mask[ i ]	
}	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id_len_minus1[ j ]	u(3)
vps_nuh_layer_id_present_flag	u(1)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( lsIdx = 1; lsIdx <= vps_num_layer_sets_minus1; lsIdx ++ ) {	
vps_profile_present_flag[ lsIdx ]	u(1)
if( !vps_profile_present_flag[ lsIdx ] )	
profile_layer_set_ref_minus1[ lsIdx ]	ue(v)
profile_tier_level( vps_profile_present_flag[ lsIdx ], vps_max_sub_layers_minus1 )	
}	
num_output_layer_sets	ue(v)
for( i = 0; i < num_output_layer_sets; i++ ) {	
output_layer_set_idx[ i ]	ue(v)
lsIdx = output_layer_set_idx[ i ]	
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[ lsIdx ][ j ] )	
output_layer_flag[ lsIdx ][ j ]	u(1)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
for( j = 0; j < i; j++ )	
direct_dependency_flag[ i ][ j ]	u(1)
}	

FIG. 35



[Fig. 36]

output_layer_sets_change( payloadSize ) {	Descriptor
active_vps_id	u(4)
num_changed_output_layer_sets	ue(v)
for( i = 0; i < num_changed_output_layer_sets; i++ ) {	
changed_output_layer_set_idx_entry[ i ]	ue(v)
clsIdx = output_layer_set_idx[ changed_output_layer_set_idx_entry[ i ] ]	
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[ clsIdx ][ j ] )	
output_layer_flag[ clsIdx ][ j ]	u(1)
}	
num_addl_output_layer_sets	
for( i = num_output_layer_sets; i < num_output_layer_sets + num_addl_output_layer_sets; i++ ) {	
addlIdx = addl_output_layer_sets_idx[ i ]	ue(v)
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[ addlIdx ][ j ] )	
output_layer_flag[ addlIdx ][ j ]	u(1)
}	
}	

FIG. 36

[Fig. 37]

output_layer_sets_change( payloadSize ) {	Descriptor
active_vps_id	u(4)
num_deleted_output_layer_sets	ue(v)
for( i = 0; i < num_deleted_output_layer_sets; i++ )	
deleted_output_layer_sets_idx_entry[ i ]	ue(v)
num_changed_output_layer_sets	ue(v)
for( i = 0; i < num_changed_output_layer_sets; i++ ) {	
changed_output_layer_set_idx_entry[ i ]	ue(v)
clsIdx = output_layer_set_idx[ changed_output_layer_set_idx_entry[ i ] ]	
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[ clsIdx ][ j ] )	
output_layer_flag[ clsIdx ][ j ]	u(1)
}	
num_addl_output_layer_sets	
for( i = num_output_layer_sets - num_deleted_output_layer_sets; i < num_output_layer_sets - num_deleted_output_layer_sets + num_addl_output_layer_sets; i++ ) {	
addlIdx = addl_output_layer_sets_idx[ i ]	ue(v)
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[ addlIdx ][ j ] )	
output_layer_flag[ addlIdx ][ j ]	u(1)
}	
}	

FIG. 37

[Fig. 38A]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_layer_flag	u(1)
splitting_flag	u(1)
for( i = 0, NumScalabilityTypes = 0; i < 16; i++ ) {	
scalability_mask[ i ]	u(1)
NumScalabilityTypes += scalability_mask[ i ]	
}	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id_len_minus1[ j ]	u(3)
...	
num_output_layer_sets	ue(v)
for( i = 0; i < num_output_layer_sets; i++ ) {	
output_layer_set_idx[ i ]	ue(v)
lsIdx = output_layer_set_idx[ i ]	
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[ lsIdx ][ j ] )	
output_layer_flag[ lsIdx ][ j ]	u(1)
}	
num_op_dpb_info_parameters	ue(v)
for( i = 0; i < num_op_dpb_info_parameters; i++ ) {	
operation_point_layer_set_idx[ i ]	ue(v)
for( j = 0; j <= vps_max_layer_id; j++ )	
op_dpb_info_parameters( j )	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
for( j = 0; j < i; j++ )	
direct_dependency_flag[ i ][ j ]	u(1)
}	

FIG. 38A

[Fig. 38B]

vps_extension() {	Descriptor
while(!byte_aligned())	
vps_extension_byte_aligment_reserved_one_bit	u(1)
ave_base_layer_flag	u(1)
splitting_flag	u(1)
for( i = 0, NumScalabilityTypes = 0; i < 16; i++) {	
scalability_mask[ i ]	u(1)
NumScalabilityTypes += scalability_mask[ i ]	
}	
for( j = 0; j < NumScalabilityTypes; j++)	
dimension_id_len_minus1[ j ]	u(3)
...	
num_output_layer_sets	ue(v)
for( i = 0; i < num_output_layer_sets; i++) {	
output_layer_set_idx[ i ]	ue(v)
lsIdx = output_layer_set_idx[ i ]	
for( j = 0; j <= vps_max_layer_id; j++)	
if( layer_id_included_flag[ lsIdx ][ j ] )	
output_layer_flag[ lsIdx ][ j ]	u(1)
}	
num_op_dpb_info_parameters	ue(v)
for( i = 0; i < num_op_dpb_info_parameters; i++) {	
operation_point_layer_set_idx[ i ]	ue(v)
for( j = 0; j <= vps_max_layer_id; j++)	
op_dpb_info_parameters( i, j )	
}	
for( i = 1; i <= vps_max_layers_minus1; i++) {	
for( j = 0; j < i; j++)	
direct_dependency_flag[ i ][ j ]	u(1)
}	

FIG. 38B

[Fig. 39A]

op_dpb_info_parameters( j ) {	
vps_max_sub_layers_minus1[ j ]	u(3)
vps_sub_layer_ordering_info_present_flag[ j ]	u(1)
for( k = ( vps_sub_layer_ordering_info_present_flag[ j ] ? 0 : vps_max_sub_layers_minus1[ j ] );	
k <= vps_max_sub_layers_minus1[ j ]; k++) {	
vps_max_dec_pic_buffering_minus1[ j ][ k ]	ue(v)
vps_max_num_reorder_pics[ j ][ k ]	ue(v)
vps_max_latency_increase_plus1[ j ][ k ]	ue(v)
}	
}	

FIG. 39A

[Fig. 39B]

op_dpb_info_parameters(id, j) {	
vps_max_sub_layers_minus1[id][j]	u(3)
vps_sub_layer_ordering_info_present_flag[id][j]	u(1)
for( k = ( vps_sub_layer_ordering_info_present_flag[id][j] ? 0 : vps_max_sub_layers_minus1[id][j] );	
k <= vps_max_sub_layers_minus1[id][j]; k++ ) {	
vps_max_dec_pic_buffering_minus1[id][j][k]	ue(v)
vps_max_num_reorder_pics[id][j][k]	ue(v)
vps_max_latency_increase_plus1[id][j][k]	ue(v)
}	
}	

FIG. 39B

[Fig. 40]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_layer_flag	u(1)
splitting_flag	u(1)
for( i = 0; NumScalabilityTypes = 0; i < 16; i++ ) {	
scalability_mask[i]	u(1)
NumScalabilityTypes += scalability_mask[i]	
}	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id_len_minus1[j]	u(3)
...	
num_output_layer_sets	ue(v)
for( i = 0; i < num_output_layer_sets; i++ ) {	
output_layer_set_idx[i]	ue(v)
lsIdx = output_layer_set_idx[i]	
for( j = 0; j <= vps_max_layer_id; j++ )	
if( layer_id_included_flag[lsIdx][j] )	
output_layer_flag[lsIdx][j]	u(1)
}	
num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
output_point_layer_set_idx[i]	ue(v)
oplsIdx = output_point_layer_set_idx[i]	
for( j = 0; j <= vps_max_layer_id; j++ )	
oop_dpb_info_parameters(j)	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
for( j = 0; j < i; j++ )	
direct_dependency_flag[i][j]	u(1)
}	

FIG. 40

[Fig. 41]

oop_dpb_info_parameters(j) {	
vps_max_sub_layers_minus1[j]	u(3)
vps_sub_layer_ordering_info_present_flag[j]	u(1)
for( k = ( vps_sub_layer_ordering_info_present_flag[j] ? 0 : vps_max_sub_layers_minus1[j] );	
k <= vps_max_sub_layers_minus1[j]; k++ ) {	
vps_max_dec_pic_buffering_minus1[j][k]	ue(v)
vps_max_num_reorder_pics[j][k]	ue(v)
vps_max_latency_increase_plus1[j][k]	ue(v)
}	
}	

FIG. 41

[Fig. 42]

oop_dpb_info_parameters() {	
vps_max_sub_layers_minus1	u(3)
vps_sub_layer_ordering_info_present_flag	u(1)
for( k = ( vps_sub_layer_ordering_info_present_flag ? 0 : vps_max_sub_layers_minus1 );	
k <= vps_max_sub_layers_minus1; k++ ) {	
vps_max_dec_pic_buffering_minus1[k]	ue(v)
vps_max_num_reorder_pics[k]	ue(v)
vps_max_latency_increase_plus1[k]	ue(v)
}	
}	

FIG. 42

[Fig. 43]

num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
output_point_layer_set_idx[i]	ue(v)
oplsIdx=output_point_layer_set_idx[i]	
for( j = 0; j <= vps_max_layer_id; j++ )	
oop_dpb_info_parameters()	
}	

FIG. 43

[Fig. 44]

oop_dpb_info_parameters(id, j) {	
vps_max_sub_layers_minus1[id][j]	u(3)
vps_sub_layer_ordering_info_present_flag[id][j]	u(1)
for( k = ( vps_sub_layer_ordering_info_present_flag[id][j] ? 0 : vps_max_sub_layers_minus1[id][j] ); k <= vps_max_sub_layers_minus1[id][j]; k++ ) {	
vps_max_dec_pic_buffering_minus1[id][j][k]	ue(v)
vps_max_num_reorder_pics[id][j][k]	ue(v)
vps_max_latency_increase_plus1[id][j][k]	ue(v)
}	
}	

FIG. 44

[Fig. 45]

num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
output_point_layer_set_idx[i]	ue(v)
oplsIdx=output_point_layer_set_idx[i]	
for( j = 0; j <= vps_max_layer_id; j++ )	
oop_dpb_info_parameters(oplsIdx, j)	
}	

FIG. 45

[Fig. 46]

num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
output_point_layer_set_idx[i]	ue(v)
oplsIdx=output_point_layer_set_idx[i]	
for( j = 0; j <= vps_max_layer_id; j++ )	
oop_dpb_info_parameters(i, j)	
}	

FIG. 46

[Fig. 47]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_layer_flag	u(1)
splitting_flag	u(1)
for( i = 0, NumScalabilityTypes = 0; i < 16; i++ ) {	
scalability_mask[ i ]	u(1)
NumScalabilityTypes += scalability_mask[ i ]	
}	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id_len_minus1[ j ]	u(3)
...	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
for( j = 0; j < i; j++ )	
direct_dependency_flag[ i ][ j ]	u(1)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
layer_dpb_info( i )	
}	
}	
}	

layer_dpb_info( i ) {	
vps_max_sub_layers_minus1[ i ]	u(3)
vps_sub_layer_ordering_info_present_flag[ i ]	u(1)
for( k = ( vps_sub_layer_ordering_info_present_flag[ i ] ? 0 : vps_max_sub_layers_minus1[ i ] ); k <= vps_max_sub_layers_minus1[ i ]; k++ ) {	
vps_max_dec_pic_buffering_minus1[ i ][ k ]	ue(v)
vps_max_num_reorder_pics[ i ][ k ]	ue(v)
vps_max_latency_increase_plus1[ i ][ k ]	ue(v)
}	
}	

FIG. 47

[Fig. 48]

oop_dpb_info_parameters(id) / op_dpb_info_parameters( id )/ layer_dpb_info( id ) {	
vps_sub_layer_ordering_info_present_flag[ id ]	u(1)
for( i = ( vps_sub_layer_ordering_info_present_flag ? 0 : vps_max_sub_layers_minus1 ); i <= vps_max_sub_layers_minus1; i++ ) {	
vps_max_dec_pic_buffering_minus1[ id ][ i ]	ue(v)
vps_max_num_reorder_pics[ id ][ i ]	ue(v)
vps_max_latency_increase_plus1[ id ][ i ]	ue(v)
}	
}	

FIG. 48

[Fig. 49A]

vps_extension() {	Descriptor
while( !byte_aligned( ) )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
avc_base_layer_flag	u(1)
splitting_flag	u(1)
for( i = 0, NumScalabilityTypes = 0; i < 16; i++ ) {	
scalability_mask[ i ]	u(1)
NumScalabilityTypes += scalability_mask[ i ]	
}	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id_len_minus1[ j ]	u(3)
vps_nuh_layer_id_present_flag	u(1)
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( vps_nuh_layer_id_present_flag )	
layer_id_in_nuh[ i ]	u(6)
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
max_sub_layers_vps_predict_flag[ i ]	u(1)
if( !max_sub_layers_vps_predict_flag[ i ] )	
max_sub_layers_vps_minus1[ i ]	u(3)
}	
vps_number_layer_sets_minus1	u(10)
vps_num_profile_tier_level_minus1	u(6)
for( i = 1; i <= vps_num_profile_tier_level_minus1; i++ ) {	
vps_profile_present_flag[ i ]	u(1)
if( !vps_profile_present_flag[ i ] )	
profile_ref_minus1[ i ]	u(6)
profile_tier_level( vps_profile_present_flag[ i ], vps_max_sub_layers_minus1 )	
}	

FIG. 49A



[Fig. 49B]

multiple_output_layer_sets_in_layer_setmore_output_layer_sets_than_default_flag	u(1)
if( !multiple_output_layer_sets_in_layer_setmore_output_layer_sets_than_default_flag )	
numOutputLayerSets = vps_number_layer_sets_minus1 + 1	
else {	
num_add_output_layer_sets	u(10)
numOutputLayerSets = num_add_output_layer_sets + vps_number_layer_sets_minus1 + 1	
}	
if( numOutputLayerSets > 1 )	
default_one_target_output_layer_flag	u(1)
for( i = 1; i < numOutputLayerSets; i++ ) {	
if( i > vps_num_layer_sets_minus1 ) {	
output_layer_set_idx_minus1[ i ]	u(v)
lsIdx = output_layer_set_idx_minus1[ i ] + 1	
for( j = 0; j < NumLayersInIdList[ lsIdx ] - 1; j++ )	
output_layer_flag[ i ][ j ]	u(1)
}	
profile_level_tier_idx[ i ]	u(v)
}	
num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
output_point_layer_set_idx[ i ]	ue(v)
oop_dpb_maxbuffering_parameters( i )	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
layer_dpb_info_parameters( i )	
}	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ )	
for( j = 0; j < i; j++ )	
direct_dependency_flag[ i ][ j ]	u(1)
}	

FIG. 49B

[Fig. 50]

oop_dpb_maxbuffering_parameters( i ) {	
sub_layer_vps_buf_info_present_flag[ i ]	u(1)
for( j = (sub_layer_vps_buf_info_present_flag[ i ] ? 0 : MaxSubLayers[ i ] - 1); j <= MaxSubLayers[ i ] - 1; j++ )	
max_vps_dec_pic_buffering_minus1[ i ][ j ]	ue(v)
}	

FIG. 50

[Fig. 51]

layer_dpb_info_parameters(i) {	
sub_layer_vps_ordering_info_present_flag[i]	u(1)
for( j = (sub_layer_vps_ordering_info_present_flag[i] ? 0 : max_sub_layers_vps_minus1[i]); j <= max_sub_layers_vps_minus1[i]; j++ ) {	
max_vps_num_reorder_pics[i][j]	ue(v)
max_vps_latency_increase_plus1[i][j]	ue(v)
}	
}	

FIG. 51

[Fig. 52]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
...	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[i][j]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
max_sub_layers_vps_predict_flag[i]	u(1)
if( !max_sub_layers_vps_predict_flag[i] )	
max_sub_layers_vps_minus1[i]	u(3)
}	
...	
profile_level_tier_idx[i]	u(v)
}	
num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
if( i>0 && ( MaxSubLayers[i] == MaxSubLayers[i-1] ) )	
sub_layer_vps_buf_info_predict_flag[i]	u(1)
if( !sub_layer_vps_buf_info_predict_flag[i] ) {	
output_point_layer_set_idx[i]	ue(v)
oop_dpb_maxbuffering_parameters(i)	
}	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( max_sub_layers_vps_predict_flag[i] )	
sub_layer_vps_ordering_info_predict_flag[i]	u(1)
if( !sub_layer_ordering_info_predict_flag[i] )	
layer_dpb_info_parameters(i)	
}	
}	
...	
}	

FIG. 52

[Fig. 53]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
...	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
max_sub_layers_vps_predict_flag[ i ]	u(1)
if( !max_sub_layers_vps_predict_flag[ i ] )	
max_sub_layers_vps_minus1[ i ]	u(3)
}	
...	
profile_level_tier_idx[ i ]	u(v)
}	
num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
output_point_layer_set_idx[ i ]	ue(v)
for( k = 1; k <= vps_max_layers_minus1; i++ ) {	
oop_dpb_maxbuffering_parameters( i, k )	
}	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
layer_dpb_info_parameters( i )	
}	
}	
...	
}	

FIG. 53

[Fig. 54]

oop_dpb_maxbuffering_parameters( i, k ) {	
sub_layer_vps_buf_info_present_flag[ i ][ k ]	u(1)
for( j = (sub_layer_vps_buf_info_present_flag[ i ][ k ] ? 0 : max_sub_layers_vps_minus1[ k ] ); j <= max_sub_layers_vps_minus1[ k ]; j++ ); j++ )	
max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]	ue(v)
}	

FIG. 54

[Fig. 55]

oop_dpb_maxbuffering_parameters ( i, k ) {	
sub_layer_vps_buf_info_present_flag[ i ][ k ]	u(1)
for( j = (sub_layer_vps_buf_info_present_flag[ i ] ? 0 : MaxSubLayers[ i ][ k ] - 1); j <= MaxSubLayers[ i ][ k ] - 1; j++)	
max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]	ue(v)
}	

FIG. 55

[Fig. 56]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
...	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
max_sub_layers_vps_predict_flag[ i ]	u(1)
if( !max_sub_layers_vps_predict_flag[ i ] )	
max_sub_layers_vps_minus1[ i ]	u(3)
}	
...	
profile_level_tier_idx[ i ]	u(v)
}	
num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
if( i > 0 && ( MaxSubLayers[ i ] == MaxSubLayers[ i - 1 ] ) )	
sub_layer_vps_buf_info_predict_flag[ i ]	u(1)
if( !sub_layer_vps_buf_info_predict_flag[ i ] ) {	
output_point_layer_set_idx[ i ]	ue(v)
for( k = 1; k <= vps_max_layers_minus1; i++ ) {	
oop_dpb_maxbuffering_parameters( i, k )	
}	
}	
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( max_sub_layers_vps_predict_flag[ i ] )	
sub_layer_vps_ordering_info_predict_flag[ i ]	u(1)
if( !sub_layer_ordering_info_predict_flag[ i ] )	
layer_dpb_info_parameters( i )	
}	
}	
...	
}	

FIG. 56

[Fig. 57]

oop_dpb_maxbuffering_parameters (i, k) {	
sub_layer_vps_buf_info_present_flag[i][k]	u(1)
for( j = (sub_layer_vps_buf_info_present_flag[i][k] ? 0 : max_sub_layers_vps_minus1[k]; j <= max_sub_layers_vps_minus1[k]; j++ ); j++ )	
max_vps_dec_pic_buffering_minus1[i][k][j]	ue(v)
}	

FIG. 57

[Fig. 58]

oop_dpb_maxbuffering_parameters (i, k) {	
sub_layer_vps_buf_info_present_flag[i][k]	u(1)
for( j = (sub_layer_vps_buf_info_present_flag[i] ? 0 : MaxSubLayers[i][k] - 1); j <= MaxSubLayers[i][k] - 1; j++ )	
max_vps_dec_pic_buffering_minus1[i][k][j]	ue(v)
}	

FIG. 58

[Fig. 59]

vps_extension() {	Descriptor
while( !byte_aligned() )	
vps_extension_byte_alignment_reserved_one_bit	u(1)
...	
for( j = 0; j < NumScalabilityTypes; j++ )	
dimension_id[ i ][ j ]	u(v)
}	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
max_sub_layers_vps_predict_flag[ i ]	u(1)
if( !max_sub_layers_vps_predict_flag[ i ] )	
max_sub_layers_vps_minus1[ i ]	u(3)
}	
...	
profile_level_tier_idx[ i ]	u(v)
}	
num_dpb_info_parameters	ue(v)
for( i = 0; i < num_dpb_info_parameters; i++ ) {	
if( i > 0 && ( MaxSubLayers[ i ] == MaxSubLayers[ i - 1 ] ) )	
sub_layer_vps_buf_info_predict_flag[ i ]	u(1)
if( !sub_layer_vps_buf_info_predict_flag[ i ] ) {	
output_point_layer_set_idx[ i ]	ue(v)
for( k = 1; k <= vps_max_layers_minus1; i++ ) {	
oop_dpb_maxbuffering_parameters( i, k )	
}	
}	
}	
}	
layer_dpb_info_parameters_presence_flag	u(1)
if( layer_dpb_info_parameters_presence_flag ) {	
for( i = 1; i <= vps_max_layers_minus1; i++ ) {	
if( max_sub_layers_vps_predict_flag[ i ] )	
sub_layer_vps_ordering_info_predict_flag[ i ]	u(1)
if( !sub_layer_ordering_info_predict_flag[ i ] )	
layer_dpb_info_parameters( i )	
}	
}	
}	
...	
}	

FIG. 59

[Fig. 60]

oop_dpb_maxbuffering_parameters ( i , k ) {	
sub_layer_vps_maxbuffering_info_present_flag[ i ][ k ]	u(1)
for( j = (sub_layer_vps_maxbuffering_info_present_flag[ i ][ k ] ? 0 : max_sub_layers_vps_minus1[ k ] ); j <= max_sub_layers_vps_minus1[ k ]; j++ )	
max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]	ue(v)
}	

FIG. 60

[Fig. 61]

oop_dpb_maxbuffering_parameters ( i , k ) {	
sub_layer_vps_buf_info_present_flag[ i ][ k ]	u(1)
for( j = (sub_layer_vps_buf_info_present_flag[ i ] ? 0 : MaxSubLayers[ i ][ k ] - 1); j <= MaxSubLayers[ i ][ k ] - 1; j++ )	
max_vps_dec_pic_buffering_minus1[ i ][ k ][ j ]	ue(v)
}	

FIG. 61

[Fig. 62]

seq_parameter_set_rbsp() {	Descriptor
sps_video_parameter_set_id	u(4)
sps_max_sub_layers_minus1	u(3)
sps_temporal_id_nesting_flag	u(1)
if( nuh_layer_id == 0 )	
profile_tier_level( 1, sps_max_sub_layers_minus1 )	
sps_seq_parameter_set_id	ue(v)
...	
log2_max_pic_order_cnt_lsb_minus4	ue(v)
if( nuh_layer_id > 0 )	
sps_dpb_params_present_flag	u(1)
if( (nuh_layer_id == 0)    ( (nuh_layer_id > 0) && sps_dpb_params_present_flag) ) {	
sps_sub_layer_ordering_info_present_flag	u(1)
for( i = ( sps_sub_layer_ordering_info_present_flag ? 0 : sps_max_sub_layers_minus1 );	
i <= sps_max_sub_layers_minus1; i++ ) {	
sps_max_dec_pic_buffering_minus1[ i ]	ue(v)
sps_max_num_reorder_pics[ i ]	ue(v)
sps_max_latency_increase_plus1[ i ]	ue(v)
}	
}	
log2_min_luma_coding_block_size_minus3	ue(v)
log2_diff_max_min_luma_coding_block_size	ue(v)
log2_min_transform_block_size_minus2	ue(v)
...	
sps_extension_flag	u(1)
if( sps_extension_flag ) {	
sps_extension()	
sps_extension2_flag	u(1)
if( sps_extension2_flag )	
while( more_rbsp_data() )	
sps_extension_data_flag	u(1)
}	
rbsp_trailing_bits()	
}	

FIG. 62



[Fig. 63]

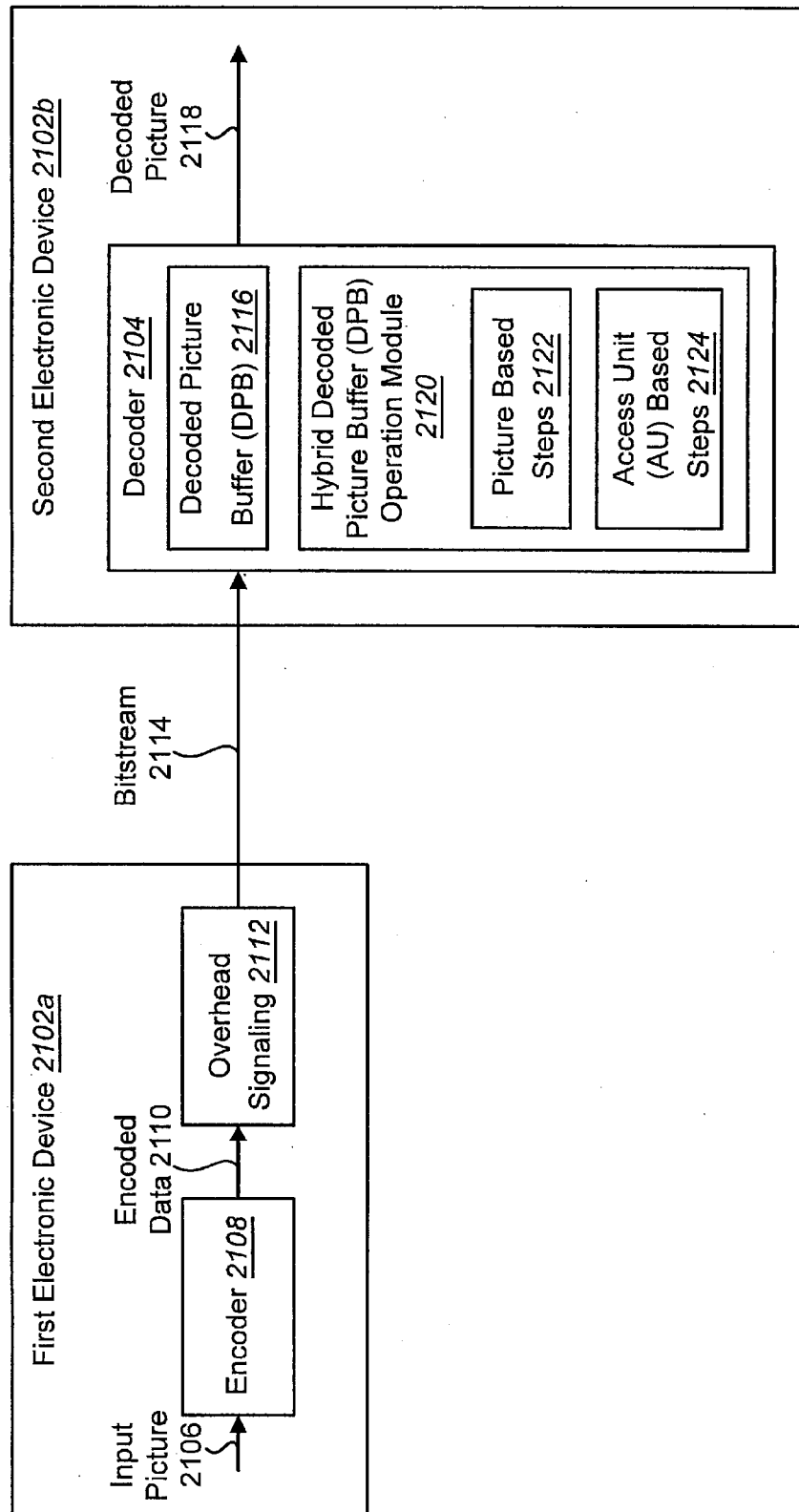
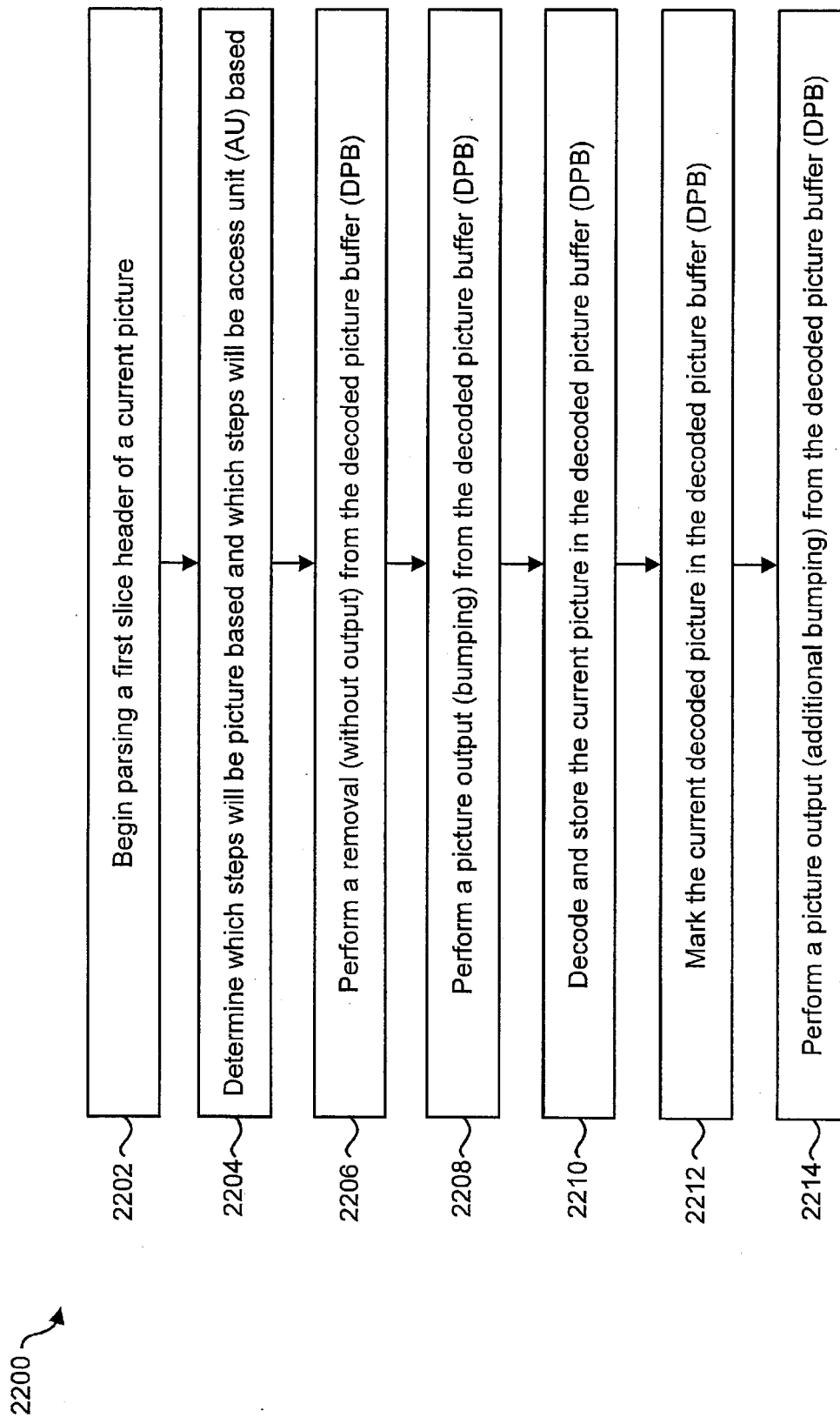
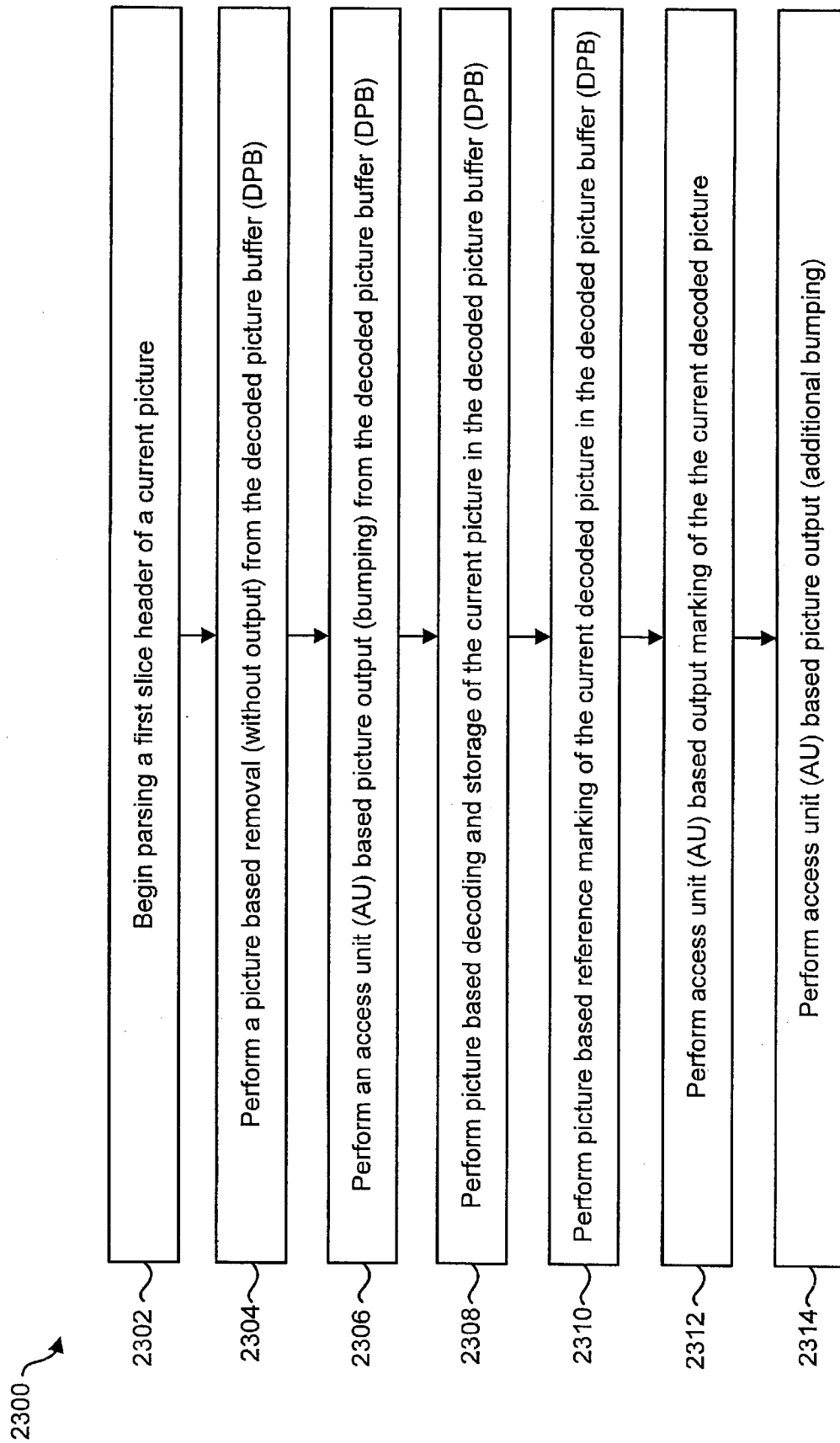


FIG. 63

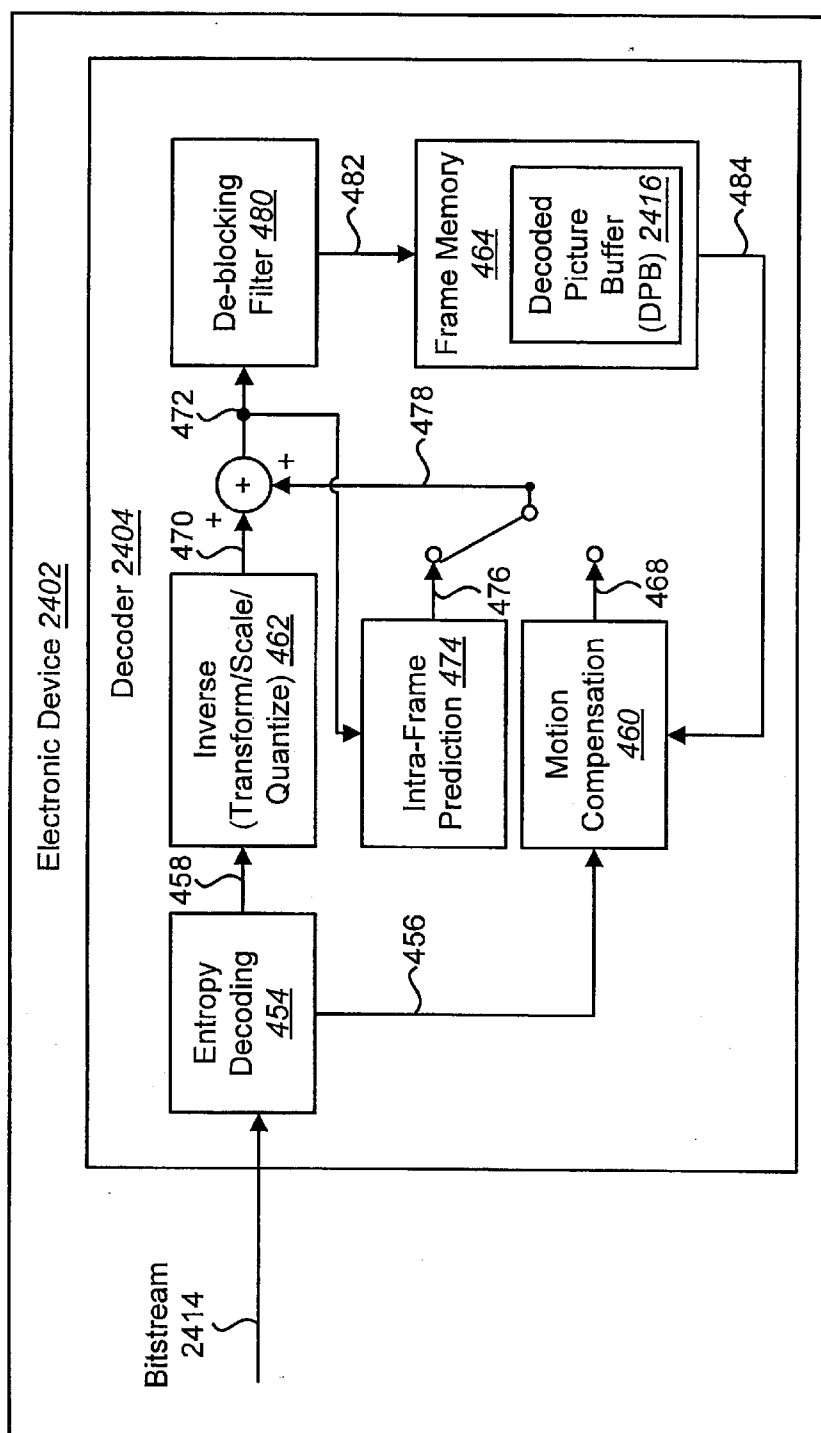
[Fig. 64]

**FIG. 64**

[Fig. 65]

**FIG. 65**

**FIG. 66**



[Fig. 67A]

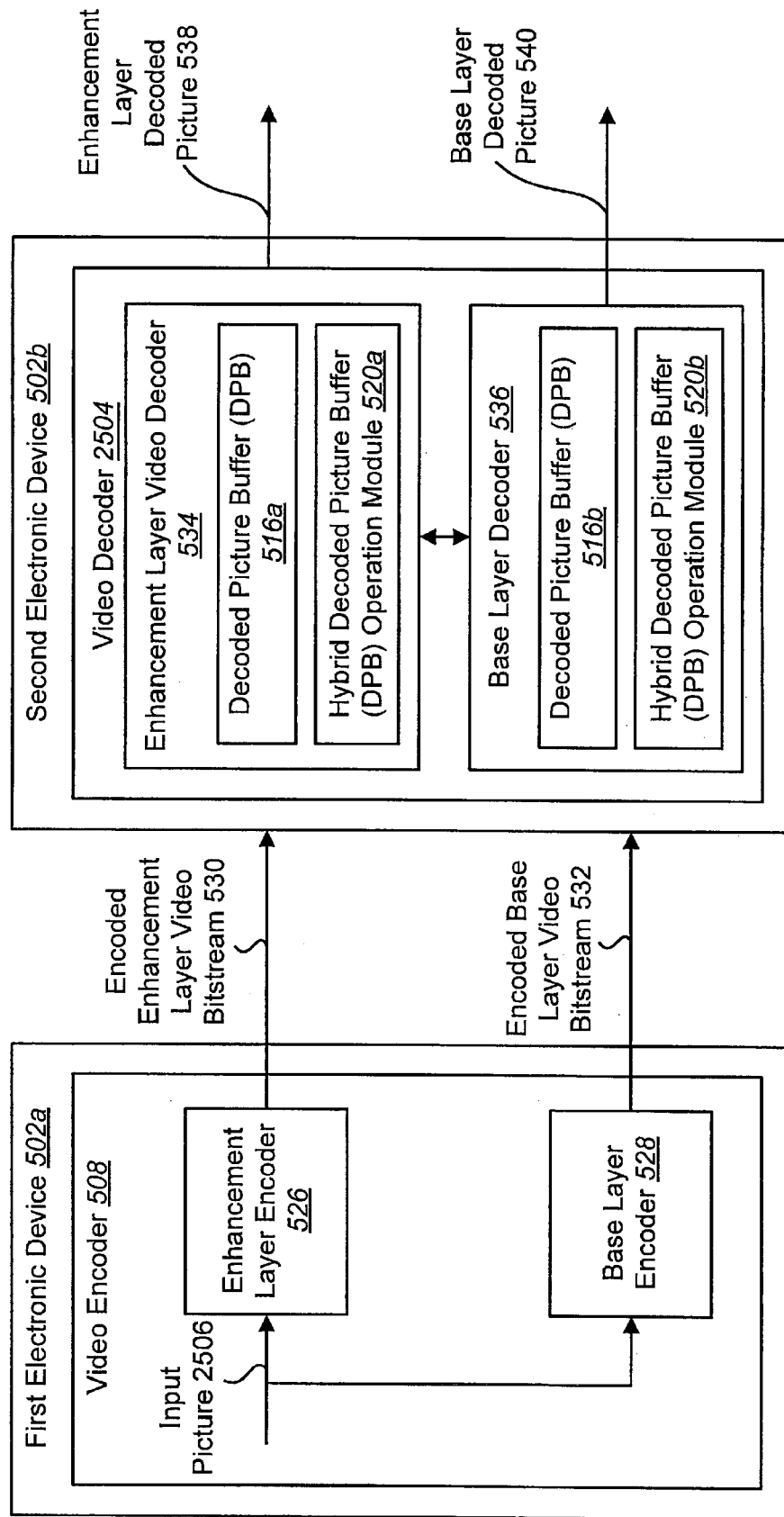


FIG. 67A

[Fig. 67B]

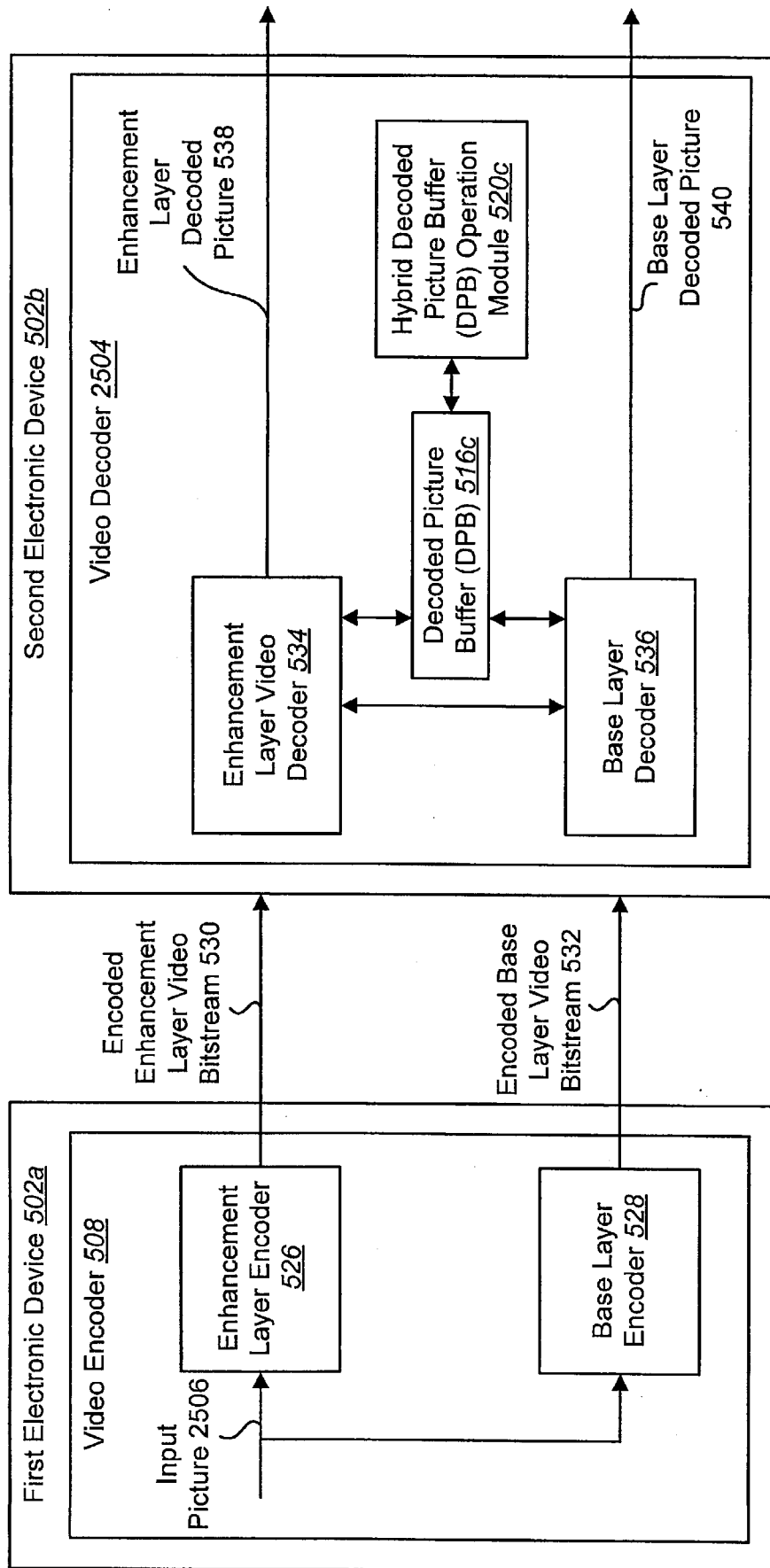


FIG. 67B

[Fig. 68]

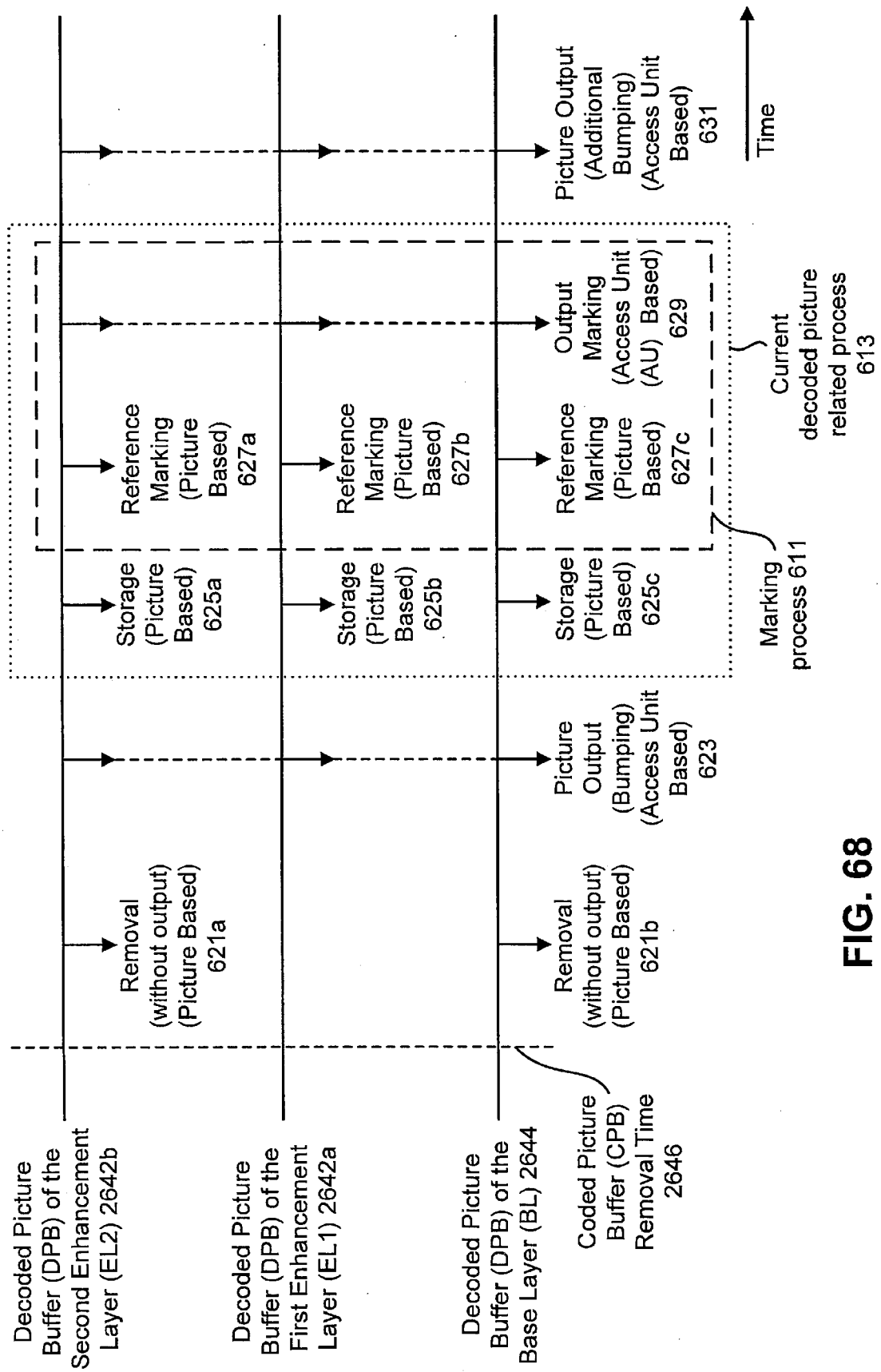


FIG. 68

[Fig. 69]

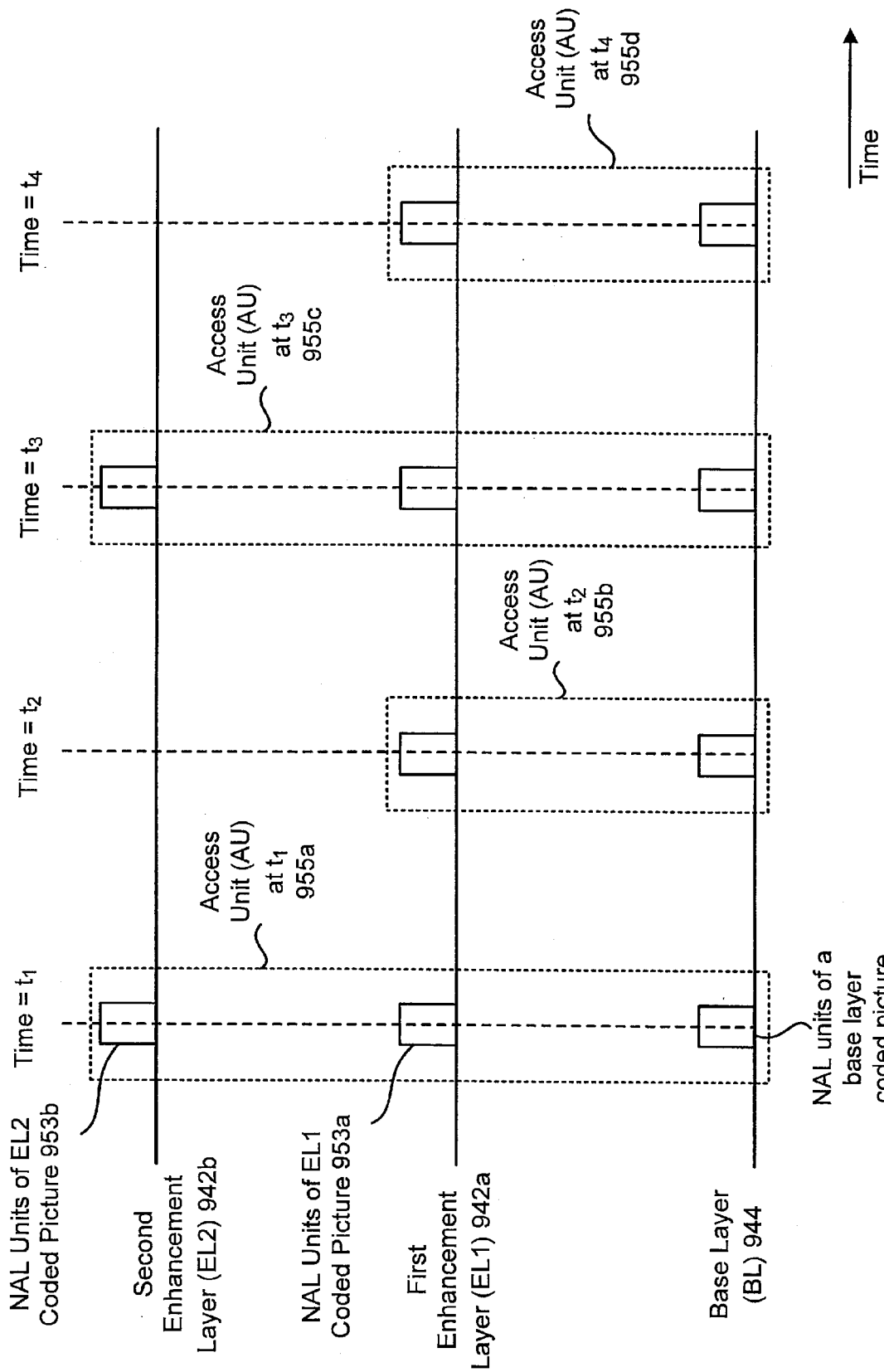


FIG. 69



[Fig. 70]

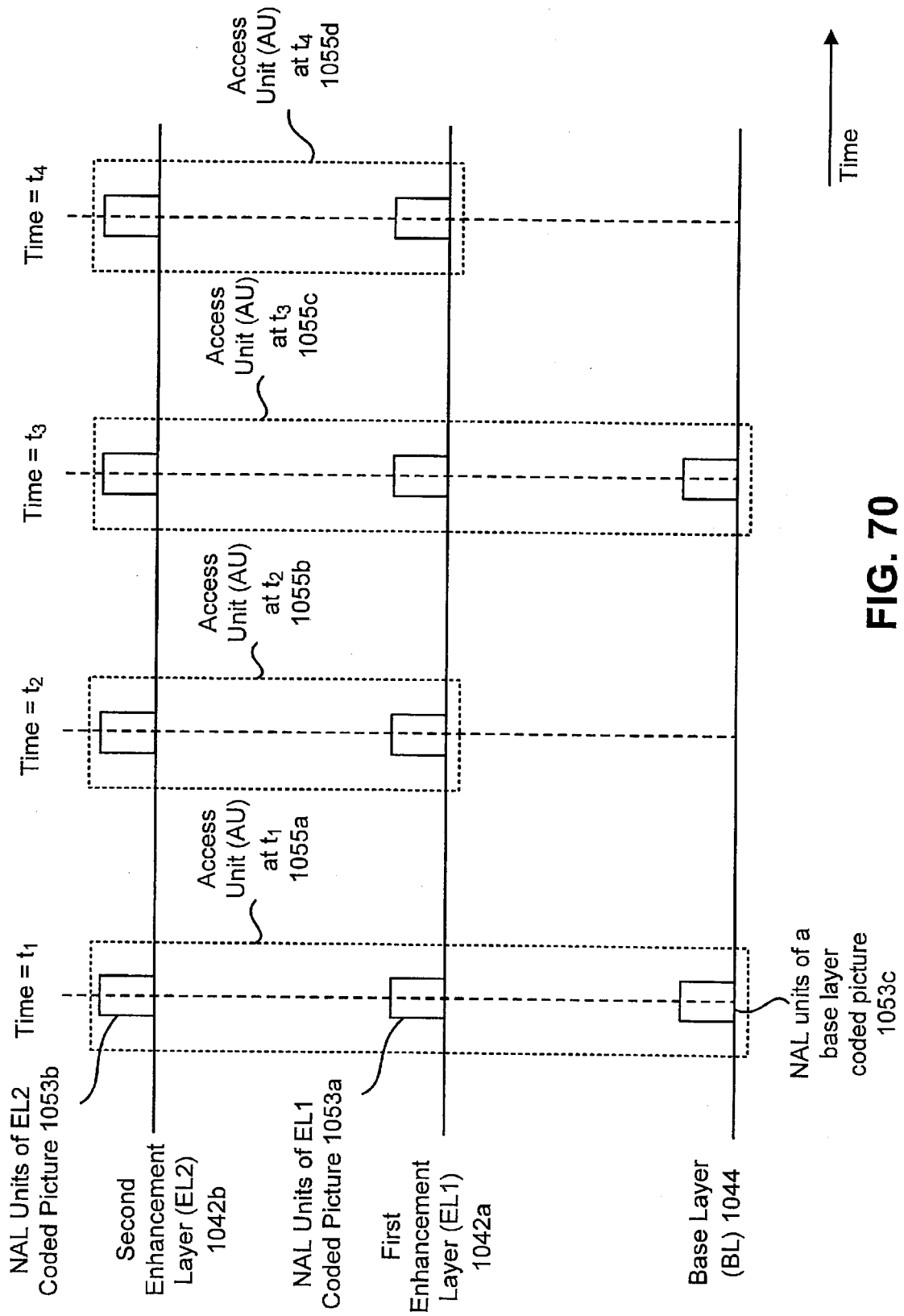


FIG. 70

