

Reference picture management methods for video coding

CROSS-REFERENCE TO RELATED APPLICATIONS

- 5 This patent application claims the priority to PCT/RU2020/000027, filed on January 23, 2020 and PCT/RU2020/000035 filed on January 27, 2020. The disclosure of the aforementioned patent applications is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

- 10 Embodiments of the present application generally relate to the field of picture processing and more particularly to video coding.

BACKGROUND

- Video coding (video encoding and decoding) is used in a wide range of digital video applications, for example broadcast digital TV, video transmission over internet and mobile networks, real-time conversational applications such as video chat, video conferencing, DVD and Blu-ray discs, video content acquisition and editing systems, and camcorders of security applications.
- 20 The amount of video data needed to depict even a relatively short video can be substantial, which may result in difficulties when the data is to be streamed or otherwise communicated across a communications network with limited bandwidth capacity. Thus, video data is generally compressed before being communicated across modern day telecommunications networks. The size of a video could also be an issue when the video is stored on a storage device because memory resources may be limited. Video compression devices often use software and/or hardware at the source to code the video data prior to transmission or storage, thereby decreasing the quantity of data needed to represent digital video images. The compressed data is then received at the destination by a video decompression device that decodes the video data. With limited network resources and ever increasing demands of
- 25 higher video quality, improved compression and decompression techniques that improve compression ratio with little to no sacrifice in picture quality are desirable.
- 30 Particularly, efficient implementation of reference picture management schemes in the context of inter prediction processing still poses a demanding problem in view of an efficient handling of and selection from different schemes that, in principle, may be made available.

SUMMARY

Embodiments of the present application provide apparatuses and methods for encoding and decoding according to the independent claims.

5

The foregoing and other objects are achieved by the subject matter of the independent claims. Further implementation forms are apparent from the dependent claims, the description and the figures.

10 Particular embodiments are outlined in the attached independent claims, with other embodiments in the dependent claims.

According to one aspect, it is provided a method of decoding an encoded video sequence, the method comprising

obtaining a value of a reference picture list (RPL) syntax element (e.g. `sps_rpl_flag`) from a
15 bitstream of the video sequence, the RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream;

when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream: a) obtaining, from the bitstream, values of the syntax elements related to reference picture lists and b)

20 constructing at least one reference picture list for inter prediction using the obtained values of the syntax elements related to reference picture lists.

In particular, it is provided a method of decoding an encoded video sequence, the method comprising

25 obtaining a value of a reference picture list (RPL) syntax element from a bitstream of the video sequence, the RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream; and

when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream, obtaining, from the bitstream, values of the syntax elements

30 related to reference picture lists.

The value of the RPL syntax element may be obtained by parsing the bitstream. Different from the art, a syntax element (RPL syntax element) is provided that allows to detect (for example,

by a decoding device) whether or not syntax related to reference picture lists is present in the bitstream and, thus, whether a reference picture lists method can be used for an inter prediction process or not. Thereby, the management of the inter prediction process, particularly, the reference picture management can be efficiently controlled.

5

According to an implementation, the method further comprises, when the RPLsyntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, constructing at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream. Construction of at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream may comprise performing a sliding window method to construct at least one reference picture list for inter prediction.

10

Thus, reference picture list management that is not based on reference picture list can be performed. In this case, a sliding window method may be performed. Accordingly, an efficient management of different reference picture management schemes including reference picture lists management as utilized by VVC, on the one hand, and a sliding window method as utilized by AVC, on the other hand, can be provided.

15

According to an implementation, the method further comprises, when the RPL syntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, obtaining, from the bitstream, a value of a second syntax element (e.g. max_num_tid0_ref_pics) related to reference picture marking (as it is utilized by AVC). The reference picture marking may provide information on a maximum number of reference pictures (in temporal layer 0) usable for inter prediction of pictures of the video sequence. This information can be readily used to determine the size of the Decoded Picture Buffer.

20

25

According to an implementation, the syntax elements related to reference picture lists are first syntax elements related to reference picture lists constructing (as it is utilized by VVC). These first syntax elements related to reference picture lists constructing may be included in at least one reference picture list syntax structure, wherein each reference picture list syntax structure is used to construct one (single) reference picture list. Employment of such a reference picture list syntax structure allows for an efficient coding of the information necessary for the construction of the reference picture lists.

30

According to an implementation, the RPL syntax element is a flag (for example, named “sps_rpl_flag”, see also detailed description below) wherein a first value (for example, 1) of the flag specifies that syntax related to reference picture lists is present in the bitstream and a
5 second value (for example, 0) of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream. Detecting/signalling of whether or not syntax related to reference picture lists is present in the bitstream by means of such a flag is very efficient in terms of the bit size of the coding.

10 In principle, the value of the RPL syntax element may be obtained from one of a sequence parameter set, a picture parameter set and a slice header. Thus, signalling on different hierarchy levels of parameter sets is envisaged.

According to an implementation, the syntax elements related to reference picture lists are
15 obtained from a sequence parameter set and/or a slice header. Again, different hierarchy levels of parameter sets may suitably be used.

According to an implementation the method of decoding comprises, when the RPL syntax element specifies that syntax related to reference picture lists is present in the bitstream,
20 obtaining a value of a at least one third syntax element (e.g. ref_pic_list_sps_flag) from a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header. Thereby, the management of the inter prediction process based on reference picture lists can be efficiently implemented.

25 According to an implementation, the at least one third syntax element comprises two third syntax elements, wherein the value of one of the two third syntax elements is obtained for a first reference picture list and, in case that a fourth syntax element (e.g. rpl1_idx_present_flag) obtained from the bitstream indicates that another third syntax element is present in the bitstream, the value of the other one of the two third syntax elements is obtained for a second
30 reference list.

The fourth syntax element may be another flag, a value of 1 of the other flag indicating that the second reference list is present in the bitstream. Employment of such a flag allows for an efficient signalling in terms of the bit size of the coding.

- According to an implementation, the at least one value of the third syntax element is obtained in case that there is at least one reference picture list syntax structure signalled at sequence parameter set level in the bitstream, wherein each reference picture list syntax structure is used
- 5 to construct one reference picture list.

The at least one value of the third syntax element may be obtained from a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

10

According to another implementation, the at least one value of the third syntax element is obtained from a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)

} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

According to a second aspect, a method of encoding a video sequence is provided corresponding to the above-described decoding method and providing the same or similar advantages as discussed above. The method of encoding a video sequence according to the second aspect comprises

5 determining whether values of syntax elements related to reference picture lists are to be used for construction (by a decoding device) of at least one reference picture list for inter prediction; generating a bitstream comprising a RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream, wherein the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream in case that it is determined that the values of the syntax elements related to reference picture lists are to be used (by the decoding device) for construction of at least one reference picture list for inter prediction.

15 According to an implementation, the bitstream comprises a second syntax element related to reference picture marking when it is determined that values of syntax elements related to reference picture lists are not to be used for construction of at least one reference picture list for inter prediction.

20 The syntax elements related to reference picture lists may be first syntax elements related to reference picture lists constructing.

The first syntax elements related to reference picture lists constructing may be included in at least one reference picture list syntax structure, wherein each reference picture list syntax structure is used for construction of one reference picture list.

25

According to an implementation, the second syntax element related to reference picture marking provides information on a maximum number of reference pictures (in temporal layer 0) usable for inter prediction of pictures of the video sequence.

- 5 According to an implementation, the RPL syntax element is a flag and a first value (for example 1) of the flag specifies that syntax related to reference picture lists is present in the bitstream and a second value (for example, 0) of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.
- 10 The RPL syntax element may be signalled in one of a sequence parameter set, picture parameter set and slice header.

The syntax elements related to reference picture lists may be comprised in a sequence parameter set and/or a slice header.

15

According to an implementation, at least one third syntax element is signalled in a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header, when it is determined that values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture

20 list for inter prediction.

According to an implementation, the bitstream comprises a fourth syntax element and the at least one third syntax element comprises two third syntax elements, wherein one of the two third syntax elements is for a first reference picture list and the other one of the two third syntax

25 elements is for a second reference list in case that a value of the fourth syntax element indicates that the other third syntax element is present in the bitstream.

The fourth syntax element may be another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

30

According to an implementation, the at least one reference picture list syntax structure is signalled at sequence parameter set level in the bitstream, wherein each reference picture list syntax structure is used for construction of one reference picture list.

The third syntax element may be signalled in a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

- 5 According to another implementation the third syntax element is signalled in a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

Furthermore, a decoder configured for performing the above-described embodiments of the inventive decoding method and an encoder configured for performing the above-described embodiments of the inventive encoding method are provided. Particularly, it is provided an decoder comprising processing circuitry for carrying out the above-described embodiments of
5 the inventive decoding method and an encoder comprising processing circuitry for carrying out the above-described embodiments of the inventive encoding method.

Further, it is provided a computer program product comprising program code for performing the methods described above when executed on a computer or a processor.

10

A decoder is provided that comprises one or more processors and a non-transitory computer-readable storage medium coupled to the processors and storing programming for execution by the processors, wherein the programming, when executed by the processors, configures the decoder to carry out the methods described above.

15

An encoder is provided that comprises one or more processors and a non-transitory computer-readable storage medium coupled to the processors and storing programming for execution by the processors, wherein the programming, when executed by the processors, configures the encoder to carry out the methods described above.

20

Further, it is provided a non-transitory computer-readable medium carrying a program code which, when executed by a computer device, causes the computer device to perform the methods described above.

25

Furthermore, the following decoding and encoding devices are provided that provide the same or similar advantages as discussed above.

According to another aspect, it is provided a decoding device for decoding an encoded video
30 sequence, comprising:

an obtaining unit configured for obtaining a value of a reference picture list (RPL) syntax element from a bitstream of the video sequence, the RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream and,

when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream, obtaining, from the bitstream, values of the syntax elements related to reference picture lists; and

5 a RPL construction unit for, when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream, constructing at least one reference picture list for inter prediction using the obtained values of the syntax elements related to reference picture lists.

10 Particularly, it is provided a decoding device for decoding an encoded video sequence, comprising an obtaining unit configured for obtaining a value of a reference picture list (RPL) syntax element from a bitstream of the video sequence, the RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream and, when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream, obtaining, from the bitstream, values of the syntax elements
15 related to reference picture lists.

According to an implementation, the RPL construction unit is configured for, when the RPL syntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, constructing at least one reference picture list for inter prediction
20 without further obtaining values of syntax elements from the bitstream.

The constructing at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream may comprise performing a sliding window method to construct at least one reference picture list for inter prediction.
25

According to an implementation, the obtaining unit is configured for, when the RPL syntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, obtaining, from the bitstream, a value of a second syntax element related to reference picture marking.
30

The syntax elements related to reference picture lists may be first syntax elements related to reference picture lists constructing.

According to an implementation, the first syntax elements related to reference picture lists constructing are included in at least one reference picture list syntax structure, and the RPL construction unit is configured for using each reference picture list syntax structure to construct one reference picture list.

5

The second syntax element related to reference picture marking may provide information on a maximum number of reference pictures usable for inter prediction of pictures of the video sequence.

- 10 The RPL syntax element may be a flag wherein a first value (for example, 1) of the flag specifies that syntax related to reference picture lists is present in the bitstream and a second value (for example, 0) of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.

- 15 According to an implementation, the obtaining unit is configured for obtaining the value of the RPS syntax element from one of a sequence parameter set, picture parameter set and slice header.

- 20 According to an implementation, the obtaining unit is configured for obtaining the syntax elements related to reference picture lists from a sequence parameter set and/or a slice header.

- 25 According to an implementation, the obtaining unit is configured for, when the RPL syntax element specifies that syntax related to reference picture lists is present in the bitstream, obtaining a value of at least one third syntax element from a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header.

- 30 According to an implementation, the at least one third syntax element comprises two third syntax elements, wherein the obtaining unit is configured for obtaining the value of one of the two third syntax elements for a first reference picture list and, in case that a fourth syntax element obtained from the bitstream indicates that another third syntax element is present in the bitstream, the value of the other one of the two third syntax elements for a second reference list.

48. The decoding device of claim 47, wherein the fourth syntax element is another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

5 According to an implementation, the obtaining unit is configured for obtaining the at least one value of the third syntax element in case that there is at least one reference picture list syntax structure signalled at sequence parameter set level in the bitstream, and wherein the RPL construction unit is configured for using each reference picture list syntax structure to construct one reference picture list.

10 According to an implementation, the obtaining unit is configured for obtaining the at least one value of the third syntax element from a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

15 Moreover, the obtaining unit may be configured for obtaining the at least one value of the third syntax element from a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 &&	
rpl1_idx_present_flag)))	

ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

According to another aspect, it is provided an encoding device for encoding a video sequence, comprising:

- 5 a determining unit for determining whether values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction;

- a bitstream generation unit configured for generating a bitstream comprising a RPL syntax element specifying whether syntax elements related to reference picture lists are present in the
- 10 bitstream, wherein the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream in case that it is determined by the determining unit that the values of the syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

- 15 The bitstream may comprise a second syntax element related to reference picture marking, when it is determined by the determining unit that values of syntax elements related to reference picture lists are not to be used for construction of at least one reference picture list for inter prediction.

- 20 The syntax elements related to reference picture lists may be first syntax elements related to reference picture lists constructing.

- According to an implementation, the first syntax elements related to reference picture lists constructing are included in at least one reference picture list syntax structure, wherein each
- 25 reference picture list syntax structure is used for construction of one reference picture list.

According to an implementation, the second syntax element related to reference picture marking provides information on a maximum number of reference pictures usable for inter prediction of pictures of the video sequence

5

According to an implementation, the RPL syntax element is a flag and a first value (for example, 1) of the flag specifies that syntax related to reference picture lists is present in the bitstream and a second value (for example, 0) of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.

10

The RPL syntax element may be comprised in one of a sequence parameter set, picture parameter set and slice header.

The syntax elements related to reference picture lists may be comprised in a sequence parameter set and/or a slice header.

15

According to an implementation, at least one third syntax element is comprised in a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header, when it is determined by the determining unit that values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

20

According to an implementation, the bitstream comprises a fourth syntax element and the at least one third syntax element comprises two third syntax elements, wherein one of the two third syntax elements is for a first reference picture list and the other one of the two third syntax elements is for a second reference list in case that a value of the fourth syntax element indicates that the other third syntax element is present in the bitstream.

25

The fourth syntax element may be another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

30

According to an implementation, at least one reference picture list syntax structure is comprised in a sequence parameter set, wherein each reference picture list syntax structure is used for construction of one reference picture list.

The third syntax element may be comprised in a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

5

The third syntax element may be comprised in a slice header comprising

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

10

Details of one or more embodiments are set forth in the accompanying drawings and the description below. Other features, objects, and advantages will be apparent from the description, drawings, and claims.

5 BRIEF DESCRIPTION OF THE DRAWINGS

In the following embodiments of the invention are described in more detail with reference to the attached figures and drawings, in which:

- FIG. 1A is a block diagram showing an example of a video coding system configured to implement embodiments of the invention;
- 10 FIG. 1B is a block diagram showing another example of a video coding system configured to implement embodiments of the invention;
- FIG. 2 is a block diagram showing an example of a video encoder configured to implement embodiments of the invention;
- FIG. 3 is a block diagram showing an example structure of a video decoder configured to
15 implement embodiments of the invention;
- FIG. 4 is a block diagram illustrating an example of an encoding apparatus or a decoding apparatus;
- FIG. 5 is a block diagram illustrating another example of an encoding apparatus or a decoding apparatus;
- 20 FIG. 6 is a block diagram showing an example structure of a content supply system 3100 which realizes a content delivery service;
- FIG. 7 is a block diagram showing a structure of an example of a terminal device;
- FIG. 8 is a flow chart illustrating a method of decoding an encoded video sequence according to an embodiment;
- 25 FIG. 9 is a flow chart illustrating a method of encoding a video sequence according to an embodiment;
- FIG. 10 is a block diagram showing a decoding device according to an embodiment;
- FIG. 11 is a block diagram showing an encoding device according to an embodiment.
- 30 In the following identical reference signs refer to identical or at least functionally equivalent features if not explicitly specified otherwise.

DETAILED DESCRIPTION OF THE EMBODIMENTS

In the following description, reference is made to the accompanying figures, which form part of the disclosure, and which show, by way of illustration, specific aspects of embodiments of the invention or specific aspects in which embodiments of the present invention may be used.

5 It is understood that embodiments of the invention may be used in other aspects and comprise structural or logical changes not depicted in the figures. The following detailed description, therefore, is not to be taken in a limiting sense, and the scope of the present invention is defined by the appended claims.

10 For instance, it is understood that a disclosure in connection with a described method may also hold true for a corresponding device or system configured to perform the method and vice versa. For example, if one or a plurality of specific method steps are described, a corresponding device may include one or a plurality of units, e.g. functional units, to perform the described one or plurality of method steps (e.g. one unit performing the one or plurality of
15 steps, or a plurality of units each performing one or more of the plurality of steps), even if such one or more units are not explicitly described or illustrated in the figures. On the other hand, for example, if a specific apparatus is described based on one or a plurality of units, e.g. functional units, a corresponding method may include one step to perform the functionality of the one or plurality of units (e.g. one step performing the functionality of the one or plurality
20 of units, or a plurality of steps each performing the functionality of one or more of the plurality of units), even if such one or plurality of steps are not explicitly described or illustrated in the figures. Further, it is understood that the features of the various exemplary embodiments and/or aspects described herein may be combined with each other, unless specifically noted otherwise.

25 Video coding typically refers to the processing of a sequence of pictures, which form the video or video sequence. Instead of the term “picture” the term “frame” or “image” may be used as synonyms in the field of video coding. Video coding (or coding in general) comprises two parts video encoding and video decoding. Video encoding is performed at the source
30 side, typically comprising processing (e.g. by compression) the original video pictures to reduce the amount of data required for representing the video pictures (for more efficient storage and/or transmission). Video decoding is performed at the destination side and typically comprises the inverse processing compared to the encoder to reconstruct the video pictures. Embodiments referring to “coding” of video pictures (or pictures in general) shall be

understood to relate to “encoding” or “decoding” of video pictures or respective video sequences. The combination of the encoding part and the decoding part is also referred to as CODEC (Coding and Decoding).

- 5 In case of lossless video coding, the original video pictures can be reconstructed, i.e. the reconstructed video pictures have the same quality as the original video pictures (assuming no transmission loss or other data loss during storage or transmission). In case of lossy video coding, further compression, e.g. by quantization, is performed, to reduce the amount of data representing the video pictures, which cannot be completely reconstructed at the decoder, i.e.
- 10 the quality of the reconstructed video pictures is lower or worse compared to the quality of the original video pictures.

- Several video coding standards belong to the group of “lossy hybrid video codecs” (i.e. combine spatial and temporal prediction in the sample domain and 2D transform coding for
- 15 applying quantization in the transform domain). Each picture of a video sequence is typically partitioned into a set of non-overlapping blocks and the coding is typically performed on a block level. In other words, at the encoder the video is typically processed, i.e. encoded, on a block (video block) level, e.g. by using spatial (intra picture) prediction and/or temporal (inter picture) prediction to generate a prediction block, subtracting the prediction block from the
- 20 current block (block currently processed/to be processed) to obtain a residual block, transforming the residual block and quantizing the residual block in the transform domain to reduce the amount of data to be transmitted (compression), whereas at the decoder the inverse processing compared to the encoder is applied to the encoded or compressed block to reconstruct the current block for representation. Furthermore, the encoder duplicates the
- 25 decoder processing loop such that both will generate identical predictions (e.g. intra- and inter predictions) and/or re-constructions for processing, i.e. coding, the subsequent blocks.

- In the following embodiments of a video coding system 10, a video encoder 20 and a video decoder 30 are described based on Figs. 1 to 3.

30

Fig. 1A is a schematic block diagram illustrating an example coding system 10, e.g. a video coding system 10 (or short coding system 10) that may utilize techniques of this present application. Video encoder 20 (or short encoder 20) and video decoder 30 (or short decoder

30) of video coding system 10 represent examples of devices that may be configured to perform techniques in accordance with various examples described in the present application.

As shown in FIG. 1A, the coding system 10 comprises a source device 12 configured to
5 provide encoded picture data 21 e.g. to a destination device 14 for decoding the encoded picture data 13.

The source device 12 comprises an encoder 20, and may additionally, i.e. optionally,
comprise a picture source 16, a pre-processor (or pre-processing unit) 18, e.g. a picture pre-
10 processor 18, and a communication interface or communication unit 22.

The picture source 16 may comprise or be any kind of picture capturing device, for example a camera for capturing a real-world picture, and/or any kind of a picture generating device, for example a computer-graphics processor for generating a computer animated picture, or any
15 kind of other device for obtaining and/or providing a real-world picture, a computer generated picture (e.g. a screen content, a virtual reality (VR) picture) and/or any combination thereof (e.g. an augmented reality (AR) picture). The picture source may be any kind of memory or storage storing any of the aforementioned pictures.

20 In distinction to the pre-processor 18 and the processing performed by the pre-processing unit 18, the picture or picture data 17 may also be referred to as raw picture or raw picture data 17.

Pre-processor 18 is configured to receive the (raw) picture data 17 and to perform pre-
25 processing on the picture data 17 to obtain a pre-processed picture 19 or pre-processed picture data 19. Pre-processing performed by the pre-processor 18 may, e.g., comprise trimming, color format conversion (e.g. from RGB to YCbCr), color correction, or de-noising. It can be understood that the pre-processing unit 18 may be optional component.

30 The video encoder 20 is configured to receive the pre-processed picture data 19 and provide encoded picture data 21 (further details will be described below, e.g., based on Fig. 2).

Communication interface 22 of the source device 12 may be configured to receive the encoded picture data 21 and to transmit the encoded picture data 21 (or any further processed

version thereof) over communication channel 13 to another device, e.g. the destination device 14 or any other device, for storage or direct reconstruction.

The destination device 14 comprises a decoder 30 (e.g. a video decoder 30), and may
5 additionally, i.e. optionally, comprise a communication interface or communication unit 28, a post-processor 32 (or post-processing unit 32) and a display device 34.

The communication interface 28 of the destination device 14 is configured receive the encoded picture data 21 (or any further processed version thereof), e.g. directly from the
10 source device 12 or from any other source, e.g. a storage device, e.g. an encoded picture data storage device, and provide the encoded picture data 21 to the decoder 30.

The communication interface 22 and the communication interface 28 may be configured to transmit or receive the encoded picture data 21 or encoded data 13 via a direct
15 communication link between the source device 12 and the destination device 14, e.g. a direct wired or wireless connection, or via any kind of network, e.g. a wired or wireless network or any combination thereof, or any kind of private and public network, or any kind of combination thereof.

20 The communication interface 22 may be, e.g., configured to package the encoded picture data 21 into an appropriate format, e.g. packets, and/or process the encoded picture data using any kind of transmission encoding or processing for transmission over a communication link or communication network.

25 The communication interface 28, forming the counterpart of the communication interface 22, may be, e.g., configured to receive the transmitted data and process the transmission data using any kind of corresponding transmission decoding or processing and/or de-packaging to obtain the encoded picture data 21.

30 Both, communication interface 22 and communication interface 28 may be configured as unidirectional communication interfaces as indicated by the arrow for the communication channel 13 in Fig. 1A pointing from the source device 12 to the destination device 14, or bi-directional communication interfaces, and may be configured, e.g. to send and receive messages, e.g. to set up a connection, to acknowledge and exchange any other information

related to the communication link and/or data transmission, e.g. encoded picture data transmission.

5 The decoder 30 is configured to receive the encoded picture data 21 and provide decoded picture data 31 or a decoded picture 31 (further details will be described below, e.g., based on Fig. 3 or Fig. 5).

The post-processor 32 of destination device 14 is configured to post-process the decoded picture data 31 (also called reconstructed picture data), e.g. the decoded picture 31, to obtain post-processed picture data 33, e.g. a post-processed picture 33. The post-processing
10 performed by the post-processing unit 32 may comprise, e.g. color format conversion (e.g. from YCbCr to RGB), color correction, trimming, or re-sampling, or any other processing, e.g. for preparing the decoded picture data 31 for display, e.g. by display device 34.

The display device 34 of the destination device 14 is configured to receive the post-processed
15 picture data 33 for displaying the picture, e.g. to a user or viewer. The display device 34 may be or comprise any kind of display for representing the reconstructed picture, e.g. an integrated or external display or monitor. The displays may, e.g. comprise liquid crystal displays (LCD), organic light emitting diodes (OLED) displays, plasma displays, projectors , micro LED displays, liquid crystal on silicon (LCoS), digital light processor (DLP) or any
20 kind of other display.

Although Fig. 1A depicts the source device 12 and the destination device 14 as separate devices, embodiments of devices may also comprise both or both functionalities, the source device 12 or corresponding functionality and the destination device 14 or corresponding
25 functionality. In such embodiments the source device 12 or corresponding functionality and the destination device 14 or corresponding functionality may be implemented using the same hardware and/or software or by separate hardware and/or software or any combination thereof.

30 As will be apparent for the skilled person based on the description, the existence and (exact) split of functionalities of the different units or functionalities within the source device 12 and/or destination device 14 as shown in Fig. 1A may vary depending on the actual device and application.

The encoder **20** (e.g. a video encoder 20) or the decoder **30** (e.g. a video decoder 30) or both encoder 20 and decoder 30 may be implemented via processing circuitry as shown in Fig. 1B, such as one or more microprocessors, digital signal processors (DSPs), application-specific integrated circuits (ASICs), field-programmable gate arrays (FPGAs), discrete logic,
5 hardware, video coding dedicated or any combinations thereof. The encoder 20 may be implemented via processing circuitry 46 to embody the various modules as discussed with respect to encoder 20 of FIG. 2 and/or any other encoder system or subsystem described herein. The decoder 30 may be implemented via processing circuitry 46 to embody the various modules as discussed with respect to decoder 30 of FIG. 3 and/or any other decoder
10 system or subsystem described herein. The processing circuitry may be configured to perform the various operations as discussed later. As shown in Fig. 5, if the techniques are implemented partially in software, a device may store instructions for the software in a suitable, non-transitory computer-readable storage medium and may execute the instructions in hardware using one or more processors to perform the techniques of this disclosure. Either
15 of video encoder **20** and video decoder **30** may be integrated as part of a combined encoder/decoder (CODEC) in a single device, for example, as shown in Fig. 1B.

Source device 12 and destination device 14 may comprise any of a wide range of devices, including any kind of handheld or stationary devices, e.g. notebook or laptop computers,
20 mobile phones, smart phones, tablets or tablet computers, cameras, desktop computers, set-top boxes, televisions, display devices, digital media players, video gaming consoles, video streaming devices (such as content services servers or content delivery servers), broadcast receiver device, broadcast transmitter device, or the like and may use no or any kind of operating system. In some cases, the source device 12 and the destination device 14 may be
25 equipped for wireless communication. Thus, the source device 12 and the destination device 14 may be wireless communication devices.

In some cases, video coding system **10** illustrated in Fig. 1A is merely an example and the techniques of the present application may apply to video coding settings (e.g., video encoding
30 or video decoding) that do not necessarily include any data communication between the encoding and decoding devices. In other examples, data is retrieved from a local memory, streamed over a network, or the like. A video encoding device may encode and store data to memory, and/or a video decoding device may retrieve and decode data from memory. In some examples, the encoding and decoding is performed by devices that do not communicate

with one another, but simply encode data to memory and/or retrieve and decode data from memory.

For convenience of description, embodiments of the invention are described herein, for example, by reference to High-Efficiency Video Coding (HEVC) or to the reference software of Versatile Video coding (VVC), the next generation video coding standard developed by the Joint Collaboration Team on Video Coding (JCT-VC) of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Motion Picture Experts Group (MPEG). One of ordinary skill in the art will understand that embodiments of the invention are not limited to HEVC or VVC.

Encoder and Encoding Method

Fig. 2 shows a schematic block diagram of an example video encoder 20 that is configured to implement the techniques of the present application. In the example of Fig. 2, the video encoder 20 comprises an input 201 (or input interface 201), a residual calculation unit 204, a transform processing unit 206, a quantization unit 208, an inverse quantization unit 210, and an inverse transform processing unit 212, a reconstruction unit 214, a loop filter unit 220, a decoded picture buffer (DPB) 230, a mode selection unit 260, an entropy encoding unit 270 and an output 272 (or output interface 272). The mode selection unit 260 may include an inter prediction unit 244, an intra prediction unit 254 and a partitioning unit 262. Inter prediction unit 244 may include a motion estimation unit and a motion compensation unit (not shown). A video encoder 20 as shown in Fig. 2 may also be referred to as hybrid video encoder or a video encoder according to a hybrid video codec.

The residual calculation unit 204, the transform processing unit 206, the quantization unit 208, the mode selection unit 260 may be referred to as forming a forward signal path of the encoder 20, whereas the inverse quantization unit 210, the inverse transform processing unit 212, the reconstruction unit 214, the buffer 216, the loop filter 220, the decoded picture buffer (DPB) 230, the inter prediction unit 244 and the intra-prediction unit 254 may be referred to as forming a backward signal path of the video encoder 20, wherein the backward signal path of the video encoder 20 corresponds to the signal path of the decoder (see video decoder 30 in Fig. 3). The inverse quantization unit 210, the inverse transform processing unit 212, the reconstruction unit 214, the loop filter 220, the decoded picture buffer (DPB) 230, the inter prediction unit 244 and the intra-prediction unit 254 are also referred to forming the “built-in decoder” of video encoder 20.

High Level Syntax (HLS)

There are several high level sets of syntax elements which are designed to transmit different high level syntax elements such as tools' control flags, video sequence specific parameters (e.g. resolution, external and internal bitdepth, etc.) and some codec routines, that can be reused, e.g. by specific pictures, slices or blocks of the picture.

In modern codecs such as VVC and EVC the following 3 parameter sets are normally used:

1) sequence parameter set (SPS); 2) picture parameter set (PPS); and 3) slice header (SH).

SPS is the most general set and includes information, which does not change for or within

- several coded pictures, e.g. video resolution, internal and external bitdepth are normally signalled in an SPS (or at SPS level). PPS includes an information, which may change from picture to picture, e.g. tiles or PPS level delta QP information. SH contains information that can change from slice to slice, as well as such picture related information that is relatively small or relevant only for certain slice or picture types. SPS, PPS and SH may be considered forming a hierarchy of high level syntax sets, wherein SPS defines the highest level, PPS a next lower level and SH a further lower level within such a hierarchy of levels. Embodiments may be configured to implement these three parameter sets and further parameter sets.

Pictures & Picture Partitioning (Pictures & Blocks)

- The encoder 20 may be configured to receive, e.g. via input 201, a picture 17 (or picture data 17), e.g. picture of a sequence of pictures forming a video or video sequence. The received picture or picture data may also be a pre-processed picture 19 (or pre-processed picture data 19). For sake of simplicity the following description refers to the picture 17. The picture 17 may also be referred to as current picture or picture to be coded (in particular in video coding to distinguish the current picture from other pictures, e.g. previously encoded and/or decoded pictures of the same video sequence, i.e. the video sequence which also comprises the current picture).

- A (digital) picture is or can be regarded as a two-dimensional array or matrix of samples with intensity values. A sample in the array may also be referred to as pixel (short form of picture element) or a pel. The number of samples in horizontal and vertical direction (or axis) of the array or picture define the size and/or resolution of the picture. For representation of color, typically three color components are employed, i.e. the picture may be represented or include three sample arrays. In RGB format or color space a picture comprises a corresponding red, green and blue sample array. However, in video coding each pixel is typically represented in

a luminance and chrominance format or color space, e.g. YCbCr, which comprises a luminance component indicated by Y (sometimes also L is used instead) and two chrominance components indicated by Cb and Cr. The luminance (or short luma) component Y represents the brightness or grey level intensity (e.g. like in a grey-scale picture), while the two chrominance (or short chroma) components Cb and Cr represent the chromaticity or color information components. Accordingly, a picture in YCbCr format comprises a luminance sample array of luminance sample values (Y), and two chrominance sample arrays of chrominance values (Cb and Cr). Pictures in RGB format may be converted or transformed into YCbCr format and vice versa, the process is also known as color transformation or conversion. If a picture is monochrome, the picture may comprise only a luminance sample array. Accordingly, a picture may be, for example, 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.

Embodiments of the video encoder 20 may comprise a picture partitioning unit (not depicted in Fig. 2) configured to partition the picture 17 into a plurality of (typically non-overlapping) picture blocks 203. These blocks may also be referred to as root blocks, macro blocks (H.264/AVC) or coding tree blocks (CTB) or coding tree units (CTU) (H.265/HEVC and VVC). The picture partitioning unit may be configured to use the same block size for all pictures of a video sequence and the corresponding grid defining the block size, or to change the block size between pictures or subsets or groups of pictures, and partition each picture into the corresponding blocks.

In further embodiments, the video encoder may be configured to receive directly a block 203 of the picture 17, e.g. one, several or all blocks forming the picture 17. The picture block 203 may also be referred to as current picture block or picture block to be coded.

Like the picture 17, the picture block 203 again is or can be regarded as a two-dimensional array or matrix of samples with intensity values (sample values), although of smaller dimension than the picture 17. In other words, the block 203 may comprise, e.g., one sample array (e.g. a luma array in case of a monochrome picture 17, or a luma or chroma array in case of a color picture) or three sample arrays (e.g. a luma and two chroma arrays in case of a color picture 17) or any other number and/or kind of arrays depending on the color format applied. The number of samples in horizontal and vertical direction (or axis) of the block 203

define the size of block 203. Accordingly, a block may, for example, be an $M \times N$ (M -column by N -row) array of samples, or an $M \times N$ array of transform coefficients.

Embodiments of the video encoder 20 as shown in Fig. 2 may be configured to encode the picture 17 block by block, e.g. the encoding and prediction is performed per block 203.

Embodiments of the video encoder 20 as shown in Fig. 2 may be further configured to partition and/or encode the picture by using slices (also referred to as video slices), wherein a picture may be partitioned into or encoded using one or more slices (typically non-overlapping), and each slice may comprise one or more blocks (e.g. CTUs) or one or more groups of blocks (e.g. tiles (H.265/HEVC and VVC) or bricks (VVC)).

Embodiments of the video encoder 20 as shown in Fig. 2 may be further configured to partition and/or encode the picture by using slices/tile groups (also referred to as video tile groups) and/or tiles (also referred to as video tiles), wherein a picture may be partitioned into or encoded using one or more slices/tile groups (typically non-overlapping), and each slice/tile group may comprise, e.g. one or more blocks (e.g. CTUs) or one or more tiles, wherein each tile, e.g. may be of rectangular shape and may comprise one or more blocks (e.g. CTUs), e.g. complete or fractional blocks.

20

Residual Calculation

The residual calculation unit 204 may be configured to calculate a residual block 205 (also referred to as residual 205) based on the picture block 203 and a prediction block 265 (further details about the prediction block 265 are provided later), e.g. by subtracting sample values of the prediction block 265 from sample values of the picture block 203, sample by sample (pixel by pixel) to obtain the residual block 205 in the sample domain.

25

Transform

The transform processing unit 206 may be configured to apply a transform, e.g. a discrete cosine transform (DCT) or discrete sine transform (DST), on the sample values of the residual block 205 to obtain transform coefficients 207 in a transform domain. The transform coefficients 207 may also be referred to as transform residual coefficients and represent the residual block 205 in the transform domain.

30

The transform processing unit 206 may be configured to apply integer approximations of DCT/DST, such as the transforms specified for H.265/HEVC. Compared to an orthogonal DCT transform, such integer approximations are typically scaled by a certain factor. In order to preserve the norm of the residual block which is processed by forward and inverse transforms, additional scaling factors are applied as part of the transform process. The scaling factors are typically chosen based on certain constraints like scaling factors being a power of two for shift operations, bit depth of the transform coefficients, tradeoff between accuracy and implementation costs, etc. Specific scaling factors are, for example, specified for the inverse transform, e.g. by inverse transform processing unit 212 (and the corresponding inverse transform, e.g. by inverse transform processing unit 312 at video decoder 30) and corresponding scaling factors for the forward transform, e.g. by transform processing unit 206, at an encoder 20 may be specified accordingly.

Embodiments of the video encoder 20 (respectively transform processing unit 206) may be configured to output transform parameters, e.g. a type of transform or transforms, e.g. directly or encoded or compressed via the entropy encoding unit 270, so that, e.g., the video decoder 30 may receive and use the transform parameters for decoding.

Quantization

The quantization unit 208 may be configured to quantize the transform coefficients 207 to obtain quantized coefficients 209, e.g. by applying scalar quantization or vector quantization. The quantized coefficients 209 may also be referred to as quantized transform coefficients 209 or quantized residual coefficients 209.

The quantization process may reduce the bit depth associated with some or all of the transform coefficients 207. For example, an n-bit transform coefficient may be rounded down to an m-bit Transform coefficient during quantization, where n is greater than m. The degree of quantization may be modified by adjusting a quantization parameter (QP). For example for scalar quantization, different scaling may be applied to achieve finer or coarser quantization. Smaller quantization step sizes correspond to finer quantization, whereas larger quantization step sizes correspond to coarser quantization. The applicable quantization step size may be indicated by a quantization parameter (QP). The quantization parameter may for example be an index to a predefined set of applicable quantization step sizes. For example, small quantization parameters may correspond to fine quantization (small quantization step sizes)

and large quantization parameters may correspond to coarse quantization (large quantization step sizes) or vice versa. The quantization may include division by a quantization step size and a corresponding and/or the inverse dequantization, e.g. by inverse quantization unit 210, may include multiplication by the quantization step size. Embodiments according to some standards, e.g. HEVC, may be configured to use a quantization parameter to determine the quantization step size. Generally, the quantization step size may be calculated based on a quantization parameter using a fixed point approximation of an equation including division. Additional scaling factors may be introduced for quantization and dequantization to restore the norm of the residual block, which might get modified because of the scaling used in the fixed point approximation of the equation for quantization step size and quantization parameter. In one example implementation, the scaling of the inverse transform and dequantization might be combined. Alternatively, customized quantization tables may be used and signaled from an encoder to a decoder, e.g. in a bitstream. The quantization is a lossy operation, wherein the loss increases with increasing quantization step sizes.

Embodiments of the video encoder 20 (respectively quantization unit 208) may be configured to output quantization parameters (QP), e.g. directly or encoded via the entropy encoding unit 270, so that, e.g., the video decoder 30 may receive and apply the quantization parameters for decoding.

Inverse Quantization

The inverse quantization unit 210 is configured to apply the inverse quantization of the quantization unit 208 on the quantized coefficients to obtain dequantized coefficients 211, e.g. by applying the inverse of the quantization scheme applied by the quantization unit 208 based on or using the same quantization step size as the quantization unit 208. The dequantized coefficients 211 may also be referred to as dequantized residual coefficients 211 and correspond - although typically not identical to the transform coefficients due to the loss by quantization - to the transform coefficients 207.

Inverse Transform

The inverse transform processing unit 212 is configured to apply the inverse transform of the transform applied by the transform processing unit 206, e.g. an inverse discrete cosine transform (DCT) or inverse discrete sine transform (DST) or other inverse transforms, to obtain a reconstructed residual block 213 (or corresponding dequantized coefficients 213)

in the sample domain. The reconstructed residual block 213 may also be referred to as transform block 213.

Reconstruction

- 5 The reconstruction unit 214 (e.g. adder or summer 214) is configured to add the transform block 213 (i.e. reconstructed residual block 213) to the prediction block 265 to obtain a reconstructed block 215 in the sample domain, e.g. by adding – sample by sample - the sample values of the reconstructed residual block 213 and the sample values of the prediction block 265.

10

Filtering

- The loop filter unit 220 (or short “loop filter” 220), is configured to filter the reconstructed block 215 to obtain a filtered block 221, or in general, to filter reconstructed samples to obtain filtered sample values. The loop filter unit is, e.g., configured to smooth pixel transitions, or otherwise improve the video quality. The loop filter unit 220 may comprise one or more loop filters such as a de-blocking filter, a sample-adaptive offset (SAO) filter or one or more other filters, e.g. an adaptive loop filter (ALF), a noise suppression filter (NSF), or any combination thereof. In an example, the loop filter unit 220 may comprise a de-blocking filter, a SAO filter and an ALF filter. The order of the filtering process may be the deblocking filter, SAO and ALF. In another example, a process called the luma mapping with chroma scaling (LMCS) (namely, the adaptive in-loop reshaper) is added. This process is performed before deblocking. In another example, the deblocking filter process may be also applied to internal sub-block edges, e.g. affine sub-blocks edges, ATMVP sub-blocks edges, sub-block transform (SBT) edges and intra sub-partition (ISP) edges. Although the loop filter unit 220 is shown in FIG. 2 as being an in loop filter, in other configurations, the loop filter unit 220 may be implemented as a post loop filter. The filtered block 221 may also be referred to as filtered reconstructed block 221.
- 15
20
25

- Embodiments of the video encoder 20 (respectively loop filter unit 220) may be configured to output loop filter parameters (such as SAO filter parameters or ALF filter parameters or LMCS parameters), e.g. directly or encoded via the entropy encoding unit 270, so that, e.g., a decoder 30 may receive and apply the same loop filter parameters or respective loop filters for decoding.
- 30

Decoded Picture Buffer

The decoded picture buffer (DPB) 230 may be a memory that stores reference pictures, or in general reference picture data, for encoding video data by video encoder 20. The DPB 230
5 may be formed by any of a variety of memory devices, such as dynamic random access memory (DRAM), including synchronous DRAM (SDRAM), magnetoresistive RAM (MRAM), resistive RAM (RRAM), or other types of memory devices. The decoded picture buffer (DPB) 230 may be configured to store one or more filtered blocks 221. The decoded picture buffer 230 may be further configured to store other previously filtered blocks, e.g.
10 previously reconstructed and filtered blocks 221, of the same current picture or of different pictures, e.g. previously reconstructed pictures, and may provide complete previously reconstructed, i.e. decoded, pictures (and corresponding reference blocks and samples) and/or a partially reconstructed current picture (and corresponding reference blocks and samples), for example for inter prediction. The decoded picture buffer (DPB) 230 may be also
15 configured to store one or more unfiltered reconstructed blocks 215, or in general unfiltered reconstructed samples, e.g. if the reconstructed block 215 is not filtered by loop filter unit 220, or any other further processed version of the reconstructed blocks or samples.

In a video codec specification, pictures are identified for multiple purposes, including for use
20 as a reference picture in inter-prediction, for output of pictures from the decoded picture buffer (DPB), for scaling of motion vectors, for weighted prediction, etc. In AVC and HEVC, pictures can be identified by picture order count (POC). In AVC and HEVC, pictures in the DPB can be marked as “used for short-term reference,” “used for long-term reference,” or “unused for reference.” Once a picture has been marked “unused for
25 reference,” the picture can no longer be used for prediction. When the picture is no longer needed for output, the picture can be removed from the DPB.

In AVC, there are two types of reference pictures, short-term and long-term. A reference picture may be marked as “unused for reference” when it becomes no longer needed for prediction reference. The conversion among these three statuses (short-term, long-term, and
30 unused for reference) is controlled by the decoded reference picture marking process. There are two alternative decoded reference picture marking mechanisms, *the implicit sliding window* process and the explicit memory management control operation (MMCO) process. The sliding window process marks a short-term reference picture as “unused for reference” when the number of reference frames is equal to a given maximum number

(max_num_ref_frames in the sequence parameter set (SPS)). The short-term reference pictures are stored in a first-in, first-out manner so that the most recently decoded short-term pictures are kept in the DPB.

The explicit MMCO process may include multiple MMCO commands. An MMCO

- 5 command may mark one or more short-term or long-term reference pictures as “unused for reference,” may mark all the pictures as “unused for reference,” or may mark the current reference picture or an existing short-term reference picture as long-term and then assign a long-term picture index to that long-term reference picture.

- 10 In AVC, the reference picture marking operations as well as the processes for output and removal of pictures from the DPB are performed after a picture has been decoded.

Embodiments of the invention may be configured to implement any of these methods and may be configured to use the implicit sliding window process, the explicit memory management control operation (MMCO), or both.

- Another method for reference pictures management is used in VVC. It is called an Reference
15 Picture Lists (RPL) method and includes the following aspects, each of which can be applied individually, and some of which can be applied in combination. 1) Reference picture marking is directly based on the two reference picture lists, namely reference picture list 0 and reference picture list 1. 1a) Information for derivation of the two reference picture lists is signaled based on syntax elements and syntax structures in the SPS, PPS, and/or the slice
20 header. 1b) Each of the two reference picture lists for a picture is signaled explicitly in a reference picture list structure. 1b.i) One or more reference picture list structures can be signaled in SPS and each of them can be referred to by an index from the slice header. 1b.ii) Each of the reference picture list 0 and 1 can be signaled directly in the slice header. 2) Information for derivation of the two reference picture lists is signaled for all types of
25 slices, i.e., B (bi-predictive), P (uni-predictive), and I (intra) slices. The term slice refers to a collection of coding tree units such as a slice in HEVC or the latest VVC WD or EVC WD; it may also refer to some other collection of coding tree units such as a tile in HEVC. 3) The two reference picture lists are generated for all types of slices, i.e., B, P, and I slices. 4)
The two reference picture lists are directly constructed without using a reference picture list
30 initialization process and a reference picture list modification process. 5) In each of the two reference picture lists, reference pictures that may be used for inter prediction of the current picture can only be referred to by a number of entries at the beginning of the list. These entries are referred to as the active entries in the list, while other entries are referred to as the inactive entries in the list. The number of the total entries and the number of the active

- entries in the list can both be derived. 6) The picture referred to by an inactive entry in a reference picture list is disallowed to be referred to by another entry in the reference picture list or any entry in the other reference picture list. 7) Long-term reference pictures are only identified by a certain number of POC LSBs, where this number may be greater than the number of POC LSBs signaled in the slice headers for derivation of POC values, and this number is indicated in the SPS. 8) Reference picture list structures are signaled only in slice headers, both short-term reference pictures and long-term reference pictures are identified by their POC LSBs, which may be represented by numbers of bits that are different from the number of bits used for representing the POC LSBs signaled in slice headers for derivation of POC values, and the numbers of bits used to represent the POC LSBs for identifying short-term reference pictures and long-term reference pictures may be different. 9) Reference picture list structures are signaled only in slice headers, no distinction is made between short-term and long-term reference pictures, all reference pictures are just named reference pictures, and reference pictures are identified by their POC LSBs, which may be represented by a number of bits that is different from number of bits used for representing the POC LSBs signaled in slice headers for derivation of POC values. Embodiments of the invention may be configured to implement any of these methods.

A general implementation of RPL method assumes that the following syntax elements are present in the bitstream.

seq_parameter_set_rbsp() {	Descriptor
...	
rpl1_same_as_rpl0_flag	u(1)
for(i = 0; i < !rpl1_same_as_rpl0_flag ? 2 : 1; i++) {	
num_ref_pic_lists_in_sps[i]	ue(v)
for(j = 0; j < num_ref_pic_lists_in_sps[i]; j++)	
ref_pic_list_struct(i, j, long_term_ref_pics_flag)	
}	
...	
}	

20

pic_parameter_set_rbsp() {	Descriptor
...	
rpl1_idx_present_flag	u(1)
...	
}	

slice_header() {	Descriptor
...	

if(nal_unit_type != IRAP_NUT) {	
...	
for(i = 0; i < 2; i++) {	
if((i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	

ref_pic_list_struct(listIdx, rplsIdx, ltrpFlag) {	Descriptor
num_strp_entries[listIdx][rplsIdx]	ue(v)
if(ltrpFlag)	
num_ltrp_entries[listIdx][rplsIdx]	ue(v)
for(i = 0; i < NumEntriesInList[listIdx][rplsIdx]; i++) {	
if(num_ltrp_entries[listIdx][rplsIdx] > 0)	
lt_ref_pic_flag[listIdx][rplsIdx][i]	u(1)
if(!lt_ref_pic_flag[listIdx][rplsIdx][i]) {	
delta_poc_st[listIdx][rplsIdx][i]	ue(v)
if(delta_poc_st[listIdx][rplsIdx][i] > 0)	
strp_entry_sign_flag[listIdx][rplsIdx][i]	u(1)
} else	
poc_lsb_lt[listIdx][rplsIdx][i]	u(v)
}	
}	

Where the semantics on newly added syntax elements related to RPL is given below.

rpl1_same_as_rpl0_flag equal to 1 specifies that the syntax structures

- 5 num_ref_pic_lists_in_sps[1] and ref_pic_list_struct(1, rplsIdx, ltrpFlag) are not present.

num_ref_pic_lists_in_sps[i] specifies the number of the ref_pic_list_struct(listIdx, rplsIdx, ltrpFlag) syntax structures with listIdx equal to i included in the SPS. The value of num_ref_pic_lists_in_sps[i] shall be in the range of 0 to 64, inclusive.

ref_pic_list_struct(listIdx, rplsIdx, ltrpFlag) syntax structure directly signalled in the slice headers of a current picture.

- 10

rpl1_idx_present_flag equal to 0 specifies that ref_pic_list_sps_flag[1] and

ref_pic_list_idx[1] are not present in slice headers. rpl1_idx_present_flag equal to 1

specifies that `ref_pic_list_sps_flag[ 1 ]` and `ref_pic_list_idx[ 1 ]` may be present in slice headers.

`ref_pic_list_sps_flag[ i ]` equal to 1 specifies that reference picture list `i` of the current picture is derived based on one of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structures with `listIdx` equal to `i` in the active SPS. `ref_pic_list_sps_flag[ i ]` equal to 0 specifies that reference picture list `i` of the current picture is derived based on the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure with `listIdx` equal to `i` that is directly included in the slice headers of the current picture. When `num_ref_pic_lists_in_sps[ i ]` is equal to 0, the value of `ref_pic_list_sps_flag[ i ]` shall be equal to 0. When `rpl1_idx_present_flag` is equal to 0 and `ref_pic_list_sps_flag[ 0 ]` is present, the value of `ref_pic_list_sps_flag[ 1 ]` is inferred to be equal to `ref_pic_list_sps_flag[ 0 ]`. When not present, the value of `ref_pic_list_sps_flag[ i ]` is inferred to be equal to 0.

`ref_pic_list_idx[ i ]` specifies the index, into the list of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structures with `listIdx` equal to `i` included in the active SPS, of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure with `listIdx` equal to `i` that is used for derivation of reference picture list `i` of the current picture. The syntax element `ref_pic_list_idx[ i ]` is represented by $\text{Ceil}(\text{Log2}(\text{num_ref_pic_lists_in_sps}[i]))$ bits. When not present, the value of `ref_pic_list_idx[ i ]` is inferred to be equal to 0. The value of `ref_pic_list_idx[ i ]` shall be in the range of 0 to `num_ref_pic_lists_in_sps[ i ] - 1`, inclusive. When `rpl1_idx_present_flag` is equal to 0 and `ref_pic_list_sps_flag[ 0 ]` is present, the value of `ref_pic_list_idx[ 1 ]` is inferred to be equal to `ref_pic_list_idx[ 0 ]`.

`num_strp_entries[ listIdx ][ rplIdx ]` specifies the number of STRP entries in the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure. When not present, the value of `num_strp_entries[ listIdx ][ rplIdx ]` is inferred to be equal to 0. The value of `num_strp_entries[ listIdx ][ rplIdx ]` shall be in the range of 0 to `sps_max_dec_pic_buffering_minus1`, inclusive.

`num_ltrp_entries[ listIdx ][ rplIdx ]` specifies the number of LTRP entries in the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure. When not present, the value of `num_ltrp_entries[ listIdx ][ rplIdx ]` is inferred to be equal to 0.

The variable `NumEntriesInList[ listIdx ][ rplIdx ]` is derived as follows:

$$\text{NumEntriesInList}[\text{listIdx}][\text{rplIdx}] = \text{num_strp_entries}[\text{listIdx}][\text{rplIdx}] + \text{num_ltrp_entries}[\text{listIdx}][\text{rplIdx}] \quad (123)$$

The value of `NumEntriesInList[ listIdx ][ rplIdx ]` shall be in the range of 0 to `sps_max_dec_pic_buffering_minus1`, inclusive.

`lt_ref_pic_flag[ listIdx ][ rplIdx ][ i ]` equal to 1 specifies that the `i`-th entry in the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure is an LTRP entry. `lt_ref_pic_flag[ listIdx ][ rplIdx ][ i ]` equal to 0 specifies that the `i`-th entry in the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure is an STRP entry. When not present, the value of `lt_ref_pic_flag[ listIdx ][ rplIdx ][ i ]` is inferred to be equal to 0.

The sum of `lt_ref_pic_flag[ listIdx ][ rplIdx ][ i ]` for all values of `i` in the range of 0 to `NumEntriesInList[ listIdx ][ rplIdx ] - 1`, inclusive, shall be equal to `num_ltrp_entries[ listIdx ][ rplIdx ]`.

delta_poc_st[listIdx][rplIdx][i], when the i-th entry is the first STRP entry in **ref_pic_list_struct**(rplIdx, ltrpFlag) syntax structure, specifies the absolute difference between the picture order count values of the current picture and the picture referred to by the i-th entry, or, when the i-th entry is an STRP entry but not the first STRP entry in the **ref_pic_list_struct**(rplIdx, ltrpFlag) syntax structure, specifies the absolute difference between the picture order count values of the pictures referred to by the i-th entry and by the previous STRP entry in the **ref_pic_list_struct**(listIdx, rplIdx, ltrpFlag) syntax structure.

The value of **delta_poc_st**[listIdx][rplIdx][i] shall be in the range of -2^{15} to $2^{15} - 1$, inclusive.

strp_entry_sign_flag[listIdx][rplIdx][i] equal to 1 specifies that i-th entry in the syntax structure **ref_pic_list_struct**(listIdx, rplIdx, ltrpFlag) has a value greater than or equal to 0. **strp_entry_sign_flag**[listIdx][rplIdx] equal to 0 specifies that the i-th entry in the syntax structure **ref_pic_list_struct**(listIdx, rplIdx, ltrpFlag) has a value less than 0. When not present, the value of **strp_entry_sign_flag**[i][j] is inferred to be equal to 1.

The list **DeltaPocSt**[listIdx][rplIdx] is derived as follows:

```

15      for( i = 0; i < NumEntriesInList[ listIdx ][ rplIdx ]; i++ ) {
          if( !lt_ref_pic_flag[ i ][ RplIdx[ i ][ j ] ] )
              DeltaPocSt[ listIdx ][ rplIdx ][ i ] = ( strp_entry_sign_flag[ listIdx ][ rplIdx ][ i ] ) ?
              delta_poc_st[ listIdx ][ rplIdx ][ i ] : 0 - delta_poc_st[ listIdx ][ rplIdx ][ i ]
          }

```

poc_lsb_lt[listIdx][rplIdx][i] specifies the value of the picture order count modulo **MaxPicOrderCntLsb** of the picture referred to by the i-th entry in the **ref_pic_list_struct**(listIdx, rplIdx, ltrpFlag) syntax structure. The length of the **poc_lsb_lt**[listIdx][rplIdx][i] syntax element is **log2_max_pic_order_cnt_lsb_minus4** + 4 bits.

Embodiments of the invention may be configured to implement any of these methods.

Mode Selection (Partitioning & Prediction)

The mode selection unit 260 comprises partitioning unit 262, inter-prediction unit 244 and intra-prediction unit 254, and is configured to receive or obtain original picture data, e.g. an

original block 203 (current block 203 of the current picture 17), and reconstructed picture data, e.g. filtered and/or unfiltered reconstructed samples or blocks of the same (current) picture and/or from one or a plurality of previously decoded pictures, e.g. from decoded picture buffer 230 or other buffers (e.g. line buffer, not shown).. The reconstructed picture data is used as reference picture data for prediction, e.g. inter-prediction or intra-prediction, to obtain a prediction block 265 or predictor 265.

Mode selection unit 260 may be configured to determine or select a partitioning for a current block prediction mode (including no partitioning) and a prediction mode (e.g. an intra or inter prediction mode) and generate a corresponding prediction block 265, which is used for the calculation of the residual block 205 and for the reconstruction of the reconstructed block 215.

Embodiments of the mode selection unit 260 may be configured to select the partitioning and the prediction mode (e.g. from those supported by or available for mode selection unit 260), which provide the best match or in other words the minimum residual (minimum residual means better compression for transmission or storage), or a minimum signaling overhead (minimum signaling overhead means better compression for transmission or storage), or which considers or balances both. The mode selection unit 260 may be configured to determine the partitioning and prediction mode based on rate distortion optimization (RDO), i.e. select the prediction mode which provides a minimum rate distortion. Terms like “best”, “minimum”, “optimum” etc. in this context do not necessarily refer to an overall “best”, “minimum”, “optimum”, etc. but may also refer to the fulfillment of a termination or selection criterion like a value exceeding or falling below a threshold or other constraints leading potentially to a “sub-optimum selection” but reducing complexity and processing time.

In other words, the partitioning unit 262 may be configured to partition a picture from a video sequence into a sequence of coding tree units (CTUs), and the CTU 203 may be further partitioned into smaller block partitions or sub-blocks (which form again blocks), e.g. iteratively using quad-tree-partitioning (QT), binary partitioning (BT) or triple-tree-partitioning (TT) or any combination thereof, and to perform, e.g., the prediction for each of the block partitions or sub-blocks, wherein the mode selection comprises the selection of the tree-structure of the partitioned block 203 and the prediction modes are applied to each of the block partitions or sub-blocks.

In the following the partitioning (e.g. by partitioning unit 260) and prediction processing (by inter-prediction unit 244 and intra-prediction unit 254) performed by an example video encoder 20 will be explained in more detail.

Partitioning

The partitioning unit 262 may be configured to partition a picture from a video sequence into a sequence of coding tree units (CTUs), and the partitioning unit 262 may partition (or split) a coding tree unit (CTU) 203 into smaller partitions, e.g. smaller blocks of square or rectangular size. For a picture that has three sample arrays, a CTU consists of an $N \times N$ block of luma samples together with two corresponding blocks of chroma samples. The maximum allowed size of the luma block in a CTU is specified to be 128×128 in the developing

versatile video coding (VVC), but it can be specified to be value rather than 128x128 in the future, for example, 256x256. The CTUs of a picture may be clustered/grouped as slices/tile groups, tiles or bricks. A tile covers a rectangular region of a picture, and a tile can be divided into one or more bricks. A brick consists of a number of CTU rows within a tile. A tile that is not partitioned into multiple bricks can be referred to as a brick. However, a brick is a true subset of a tile and is not referred to as a tile.. There are two modes of tile groups are supported in VVC, namely the raster-scan slice/tile group mode and the rectangular slice mode. In the raster-scan tile group mode, a slice/tile group contains a sequence of tiles in tile raster scan of a picture. In the rectangular slice mode, a slice contains a number of bricks of a picture that collectively form a rectangular region of the picture. The bricks within a rectangular slice are in the order of brick raster scan of the slice. These smaller blocks (which may also be referred to as sub-blocks) may be further partitioned into even smaller partitions. This is also referred to tree-partitioning or hierarchical tree-partitioning, wherein a root block, e.g. at root tree-level 0 (hierarchy-level 0, depth 0), may be recursively partitioned, e.g. partitioned into two or more blocks of a next lower tree-level, e.g. nodes at tree-level 1 (hierarchy-level 1, depth 1), wherein these blocks may be again partitioned into two or more blocks of a next lower level, e.g. tree-level 2 (hierarchy-level 2, depth 2), etc. until the partitioning is terminated, e.g. because a termination criterion is fulfilled, e.g. a maximum tree depth or minimum block size is reached. Blocks which are not further partitioned are also referred to as leaf-blocks or leaf nodes of the tree. A tree using partitioning into two partitions is referred to as binary-tree (BT), a tree using partitioning into three partitions is referred to as ternary-tree (TT), and a tree using partitioning into four partitions is referred to as quad-tree (QT).

25

For example, a coding tree unit (CTU) may be or comprise a CTB of luma samples, two corresponding CTBs of chroma samples of a picture that has three sample arrays, or a CTB of samples of a monochrome picture or a picture that is coded using three separate colour planes and syntax structures used to code the samples. Correspondingly, a coding tree block (CTB) may be an NxN block of samples for some value of N such that the division of a component into CTBs is a partitioning. A coding unit (CU) may be or comprise a coding block of luma samples, two corresponding coding blocks of chroma samples of a picture that has three sample arrays, or a coding block of samples of a monochrome picture or a picture that is coded using three separate colour planes and syntax structures used to code the samples.

30

Correspondingly a coding block (CB) may be an $M \times N$ block of samples for some values of M and N such that the division of a CTB into coding blocks is a partitioning.

5 In embodiments, e.g., according to HEVC, a coding tree unit (CTU) may be split into CUs by using a quad-tree structure denoted as coding tree. The decision whether to code a picture area using inter-picture (temporal) or intra-picture (spatial) prediction is made at the leaf CU level. Each leaf CU can be further split into one, two or four PUs according to the PU splitting type. Inside one PU, the same prediction process is applied and the relevant information is transmitted to the decoder on a PU basis. After obtaining the residual block by
10 applying the prediction process based on the PU splitting type, a leaf CU can be partitioned into transform units (TUs) according to another quadtree structure similar to the coding tree for the CU.

In embodiments, e.g., according to the latest video coding standard currently in development,
15 which is referred to as Versatile Video Coding (VVC), a combined Quad-tree nested multi-type tree using binary and ternary splits segmentation structure for example used to partition a coding tree unit. In the coding tree structure within a coding tree unit, a CU can have either a square or rectangular shape. For example, the coding tree unit (CTU) is first partitioned by a quaternary tree. Then the quaternary tree leaf nodes can be further partitioned
20 by a multi-type tree structure. There are four splitting types in multi-type tree structure, vertical binary splitting (SPLIT_BT_VER), horizontal binary splitting (SPLIT_BT_HOR), vertical ternary splitting (SPLIT_TT_VER), and horizontal ternary splitting (SPLIT_TT_HOR). The multi-type tree leaf nodes are called coding units (CUs), and unless the CU is too large for the maximum transform length, this segmentation is used for
25 prediction and transform processing without any further partitioning. This means that, in most cases, the CU, PU and TU have the same block size in the quadtree with nested multi-type tree coding block structure. The exception occurs when maximum supported transform length is smaller than the width or height of the colour component of the CU. VVC develops a unique signaling mechanism of the partition splitting information in quadtree with nested
30 multi-type tree coding tree structure. In the signalling mechanism, a coding tree unit (CTU) is treated as the root of a quaternary tree and is first partitioned by a quaternary tree structure. Each quaternary tree leaf node (when sufficiently large to allow it) is then further partitioned by a multi-type tree structure. In the multi-type tree structure, a first flag (mtt_split_cu_flag) is signalled to indicate whether the node is further partitioned; when a node is further

partitioned, a second flag (`mtt_split_cu_vertical_flag`) is signalled to indicate the splitting direction, and then a third flag (`mtt_split_cu_binary_flag`) is signalled to indicate whether the split is a binary split or a ternary split. Based on the values of `mtt_split_cu_vertical_flag` and `mtt_split_cu_binary_flag`, the multi-type tree slitting mode (`MttSplitMode`) of a CU can be derived by a decoder based on a predefined rule or a table. It should be noted, for a certain design, for example, 64×64 Luma block and 32×32 Chroma pipelining design in VVC hardware decoders, TT split is forbidden when either width or height of a luma coding block is larger than 64, as shown in Figure 6. TT split is also forbidden when either width or height of a chroma coding block is larger than 32. The pipelining design will divide a picture into Virtual pipeline data units s(VPDUs) which are defined as non-overlapping units in a picture. In hardware decoders, successive VPDUs are processed by multiple pipeline stages simultaneously. The VPDU size is roughly proportional to the buffer size in most pipeline stages, so it is important to keep the VPDU size small. In most hardware decoders, the VPDU size can be set to maximum transform block (TB) size. However, in VVC, ternary tree (TT) and binary tree (BT) partition may lead to the increasing of VPDUs size.s

In addition, it should be noted that, when a portion of a tree node block exceeds the bottom or right picture boundary, the tree node block is forced to be split until the all samples of every coded CU are located inside the picture boundaries.

As an example, the Intra Sub-Partitions (ISP) tool may divide luma intra-predicted blocks vertically or horizontally into 2 or 4 sub-partitions depending on the block size.

In one example, the mode selection unit 260 of video encoder 20 may be configured to perform any combination of the partitioning techniques described herein.

As described above, the video encoder 20 is configured to determine or select the best or an optimum prediction mode from a set of (e.g. pre-determined) prediction modes. The set of prediction modes may comprise, e.g., intra-prediction modes and/or inter-prediction modes.

Intra-Prediction

The set of intra-prediction modes may comprise 35 different intra-prediction modes, e.g. non-directional modes like DC (or mean) mode and planar mode, or directional modes, e.g. as defined in HEVC, or may comprise 67 different intra-prediction modes, e.g. non-directional modes like DC (or mean) mode and planar mode, or directional modes, e.g. as defined for VVC. As an example, several conventional angular intra prediction modes are adaptively

replaced with wide-angle intra prediction modes for the non-square blocks, e.g. as defined in VVC. As another example, to avoid division operations for DC prediction, only the longer side is used to compute the average for non-square blocks. And, the results of intra prediction of planar mode may be further modified by a position dependent intra prediction combination (PDPC) method.

The intra-prediction unit 254 is configured to use reconstructed samples of neighboring blocks of the same current picture to generate an intra-prediction block 265 according to an intra-prediction mode of the set of intra-prediction modes.

10

The intra prediction unit 254 (or in general the mode selection unit 260) is further configured to output intra-prediction parameters (or in general information indicative of the selected intra prediction mode for the block) to the entropy encoding unit 270 in form of syntax elements 266 for inclusion into the encoded picture data 21, so that, e.g., the video decoder 30 may receive and use the prediction parameters for decoding.

15

Inter-Prediction

The set of (or possible) inter-prediction modes depends on the available reference pictures (i.e. previous at least partially decoded pictures, e.g. stored in DBP 230) and other inter-prediction parameters, e.g. whether the whole reference picture or only a part, e.g. a search window area around the area of the current block, of the reference picture is used for searching for a best matching reference block, and/or e.g. whether pixel interpolation is applied, e.g. half/semi-pel, quarter-pel and/or 1/16 pel interpolation, or not.

20

Additional to the above prediction modes, skip mode, direct mode and/or other inter prediction mode may be applied.

25

For example, Extended merge prediction, the merge candidate list of such mode is constructed by including the following five types of candidates in order: Spatial MVP from spatial neighbor CUs, Temporal MVP from collocated CUs, History-based MVP from an FIFO table, Pairwise average MVP and Zero MVs. And a bilateral-matching based decoder side motion vector refinement (DMVR) may be applied to increase the accuracy of the MVs of the merge mode. Merge mode with MVD (MMVD), which comes from merge mode with motion vector differences. A MMVD flag is signaled right after sending a skip flag and

30

merge flag to specify whether MMVD mode is used for a CU. And a CU-level adaptive motion vector resolution (AMVR) scheme may be applied. AMVR allows MVD of the CU to be coded in different precision. Dependent on the prediction mode for the current CU, the MVDs of the current CU can be adaptively selected. When a CU is coded in merge mode, the
5 combined inter/intra prediction (CIIP) mode may be applied to the current CU. Weighted averaging of the inter and intra prediction signals is performed to obtain the CIIP prediction. Affine motion compensated prediction, the affine motion field of the block is described by motion information of two control point (4-parameter) or three control point motion vectors (6-parameter). Subblock-based temporal motion vector prediction (SbTMVP), which is
10 similar to the temporal motion vector prediction (TMVP) in HEVC, but predicts the motion vectors of the sub-CUs within the current CU. Bi-directional optical flow (BDOF), previously referred to as BIO, is a simpler version that requires much less computation, especially in terms of number of multiplications and the size of the multiplier. Triangle partition mode, in such a mode, a CU is split evenly into two triangle-shaped partitions, using either the
15 diagonal split or the anti-diagonal split. Besides, the bi-prediction mode is extended beyond simple averaging to allow weighted averaging of the two prediction signals.

The inter prediction unit 244 may include a motion estimation (ME) unit and a motion compensation (MC) unit (both not shown in Fig. 2). The motion estimation unit may be
20 configured to receive or obtain the picture block 203 (current picture block 203 of the current picture 17) and a decoded picture 231, or at least one or a plurality of previously reconstructed blocks, e.g. reconstructed blocks of one or a plurality of other/different previously decoded pictures 231, for motion estimation. E.g. a video sequence may comprise the current picture and the previously decoded pictures 231, or in other words, the current
25 picture and the previously decoded pictures 231 may be part of or form a sequence of pictures forming a video sequence.

The encoder 20 may, e.g., be configured to select a reference block from a plurality of reference blocks of the same or different pictures of the plurality of other pictures and
30 provide a reference picture (or reference picture index) and/or an offset (spatial offset) between the position (x, y coordinates) of the reference block and the position of the current block as inter prediction parameters to the motion estimation unit. This offset is also called motion vector (MV).

The motion compensation unit is configured to obtain, e.g. receive, an inter prediction parameter and to perform inter prediction based on or using the inter prediction parameter to obtain an inter prediction block 265. Motion compensation, performed by the motion compensation unit, may involve fetching or generating the prediction block based on the motion/block vector determined by motion estimation, possibly performing interpolations to sub-pixel precision. Interpolation filtering may generate additional pixel samples from known pixel samples, thus potentially increasing the number of candidate prediction blocks that may be used to code a picture block. Upon receiving the motion vector for the PU of the current picture block, the motion compensation unit may locate the prediction block to which the motion vector points in one of the reference picture lists.

The motion compensation unit may also generate syntax elements associated with the blocks and video slices for use by video decoder 30 in decoding the picture blocks of the video slice. In addition or as an alternative to slices and respective syntax elements, tile groups and/or tiles and respective syntax elements may be generated or used.

Entropy Coding

The entropy encoding unit 270 is configured to apply, for example, an entropy encoding algorithm or scheme (e.g. a variable length coding (VLC) scheme, an context adaptive VLC scheme (CAVLC), an arithmetic coding scheme, a binarization, a context adaptive binary arithmetic coding (CABAC), syntax-based context-adaptive binary arithmetic coding (SBAC), probability interval partitioning entropy (PIPE) coding or another entropy encoding methodology or technique) or bypass (no compression) on the quantized coefficients 209, inter prediction parameters, intra prediction parameters, loop filter parameters and/or other syntax elements to obtain encoded picture data 21 which can be output via the output 272, e.g. in the form of an encoded bitstream 21, so that, e.g., the video decoder 30 may receive and use the parameters for decoding, . The encoded bitstream 21 may be transmitted to video decoder 30, or stored in a memory for later transmission or retrieval by video decoder 30.

Other structural variations of the video encoder 20 can be used to encode the video stream. For example, a non-transform based encoder 20 can quantize the residual signal directly without the transform processing unit 206 for certain blocks or frames. In another implementation, an encoder 20 can have the quantization unit 208 and the inverse quantization unit 210 combined into a single unit.

Decoder and Decoding Method

Fig. 3 shows an example of a video decoder 30 that is configured to implement the techniques of this present application. The video decoder 30 is configured to receive encoded picture data 21 (e.g. encoded bitstream 21), e.g. encoded by encoder 20, to obtain a decoded picture 331. The encoded picture data or bitstream comprises information for decoding the encoded picture data, e.g. data that represents picture blocks of an encoded video slice (and/or tile groups or tiles) and associated syntax elements.

- 10 In the example of Fig. 3, the decoder 30 comprises an entropy decoding unit 304, an inverse quantization unit 310, an inverse transform processing unit 312, a reconstruction unit 314 (e.g. a summer 314), a loop filter 320, a decoded picture buffer (DBP) 330, a mode application unit 360, an inter prediction unit 344 and an intra prediction unit 354. Inter prediction unit 344 may be or include a motion compensation unit. Video decoder 30 may, in
15 some examples, perform a decoding pass generally reciprocal to the encoding pass described with respect to video encoder 100 from FIG. 2.

As explained with regard to the encoder 20, the inverse quantization unit 210, the inverse transform processing unit 212, the reconstruction unit 214, the loop filter 220, the decoded picture buffer (DPB) 230, the inter prediction unit 344 and the intra prediction unit 354 are also referred to as forming the “built-in decoder” of video encoder 20. Accordingly, the inverse quantization unit 310 may be identical in function to the inverse quantization unit 110, the inverse transform processing unit 312 may be identical in function to the inverse transform processing unit 212, the reconstruction unit 314 may be identical in function to reconstruction unit 214, the loop filter 320 may be identical in function to the loop filter 220, and the decoded picture buffer 330 may be identical in function to the decoded picture buffer 230. Therefore, the explanations provided for the respective units and functions of the video 20 encoder apply correspondingly to the respective units and functions of the video decoder 30.

30

Entropy Decoding

The entropy decoding unit 304 is configured to parse the bitstream 21 (or in general encoded picture data 21) and perform, for example, entropy decoding to the encoded picture data 21 to obtain, e.g., quantized coefficients 309 and/or decoded coding parameters (not shown in Fig.

3), e.g. any or all of inter prediction parameters (e.g. reference picture index and motion vector), intra prediction parameter (e.g. intra prediction mode or index), transform parameters, quantization parameters, loop filter parameters, and/or other syntax elements. Entropy decoding unit 304 maybe configured to apply the decoding algorithms or schemes
5 corresponding to the encoding schemes as described with regard to the entropy encoding unit 270 of the encoder 20. Entropy decoding unit 304 may be further configured to provide inter prediction parameters, intra prediction parameter and/or other syntax elements to the mode application unit 360 and other parameters to other units of the decoder 30. Video decoder 30 may receive the syntax elements at the video slice level and/or the video block level. In
10 addition or as an alternative to slices and respective syntax elements, tile groups and/or tiles and respective syntax elements may be received and/or used.

Inverse Quantization

The inverse quantization unit 310 may be configured to receive quantization parameters (QP)
15 (or in general information related to the inverse quantization) and quantized coefficients from the encoded picture data 21 (e.g. by parsing and/or decoding, e.g. by entropy decoding unit 304) and to apply based on the quantization parameters an inverse quantization on the decoded quantized coefficients 309 to obtain dequantized coefficients 311, which may also be referred to as transform coefficients 311. The inverse quantization process may include
20 use of a quantization parameter determined by video encoder 20 for each video block in the video slice (or tile or tile group) to determine a degree of quantization and, likewise, a degree of inverse quantization that should be applied.

Inverse Transform

25 Inverse transform processing unit 312 may be configured to receive dequantized coefficients 311, also referred to as transform coefficients 311, and to apply a transform to the dequantized coefficients 311 in order to obtain reconstructed residual blocks 213 in the sample domain. The reconstructed residual blocks 213 may also be referred to as transform blocks 313. The transform may be an inverse transform, e.g., an inverse DCT, an inverse
30 DST, an inverse integer transform, or a conceptually similar inverse transform process. The inverse transform processing unit 312 may be further configured to receive transform parameters or corresponding information from the encoded picture data 21 (e.g. by parsing and/or decoding, e.g. by entropy decoding unit 304) to determine the transform to be applied to the dequantized coefficients 311.

Reconstruction

The reconstruction unit 314 (e.g. adder or summer 314) may be configured to add the reconstructed residual block 313, to the prediction block 365 to obtain a reconstructed block 315 in the sample domain, e.g. by adding the sample values of the reconstructed residual block 313 and the sample values of the prediction block 365.

Filtering

The loop filter unit 320 (either in the coding loop or after the coding loop) is configured to filter the reconstructed block 315 to obtain a filtered block 321, e.g. to smooth pixel transitions, or otherwise improve the video quality. The loop filter unit 320 may comprise one or more loop filters such as a de-blocking filter, a sample-adaptive offset (SAO) filter or one or more other filters, e.g. an adaptive loop filter (ALF), a noise suppression filter (NSF), or any combination thereof. In an example, the loop filter unit 220 may comprise a de-blocking filter, a SAO filter and an ALF filter. The order of the filtering process may be the deblocking filter, SAO and ALF. In another example, a process called the luma mapping with chroma scaling (LMCS) (namely, the adaptive in-loop reshapener) is added. This process is performed before deblocking. In another example, the deblocking filter process may be also applied to internal sub-block edges, e.g. affine sub-blocks edges, ATMVP sub-blocks edges, sub-block transform (SBT) edges and intra sub-partition (ISP) edges. Although the loop filter unit 320 is shown in FIG. 3 as being an in loop filter, in other configurations, the loop filter unit 320 may be implemented as a post loop filter.

Decoded Picture Buffer

The decoded video blocks 321 of a picture are then stored in decoded picture buffer 330, which stores the decoded pictures 331 as reference pictures for subsequent motion compensation for other pictures and/or for output respectively display.

The decoder 30 is configured to output the decoded picture 311, e.g. via output 312, for presentation or viewing to a user.

Prediction

The inter prediction unit 344 may be identical to the inter prediction unit 244 (in particular to the motion compensation unit) and the intra prediction unit 354 may be identical to the inter

prediction unit 254 in function, and performs split or partitioning decisions and prediction based on the partitioning and/or prediction parameters or respective information received from the encoded picture data 21 (e.g. by parsing and/or decoding, e.g. by entropy decoding unit 304). Mode application unit 360 may be configured to perform the prediction (intra or
5 inter prediction) per block based on reconstructed pictures, blocks or respective samples (filtered or unfiltered) to obtain the prediction block 365.

When the video slice is coded as an intra coded (I) slice, intra prediction unit 354 of mode application unit 360 is configured to generate prediction block 365 for a picture block of the
10 current video slice based on a signaled intra prediction mode and data from previously decoded blocks of the current picture. When the video picture is coded as an inter coded (i.e., B, or P) slice, inter prediction unit 344 (e.g. motion compensation unit) of mode application unit 360 is configured to produce prediction blocks 365 for a video block of the current video slice based on the motion vectors and other syntax elements received from entropy decoding
15 unit 304. For inter prediction, the prediction blocks may be produced from one of the reference pictures within one of the reference picture lists. Video decoder 30 may construct the reference frame lists, List 0 and List 1, using default construction techniques based on reference pictures stored in DPB 330. The same or similar may be applied for or by embodiments using tile groups (e.g. video tile groups) and/or tiles (e.g. video tiles) in
20 addition or alternatively to slices (e.g. video slices), e.g. a video may be coded using I, P or B tile groups and /or tiles.

Mode application unit 360 is configured to determine the prediction information for a video block of the current video slice by parsing the motion vectors or related information and other
25 syntax elements, and uses the prediction information to produce the prediction blocks for the current video block being decoded. For example, the mode application unit 360 uses some of the received syntax elements to determine a prediction mode (e.g., intra or inter prediction) used to code the video blocks of the video slice, an inter prediction slice type (e.g., B slice, P slice, or GPB slice), construction information for one or more of the reference picture lists for
30 the slice, motion vectors for each inter encoded video block of the slice, inter prediction status for each inter coded video block of the slice, and other information to decode the video blocks in the current video slice. The same or similar may be applied for or by embodiments using tile groups (e.g. video tile groups) and/or tiles (e.g. video tiles) in addition or

alternatively to slices (e.g. video slices), e.g. a video may be coded using I, P or B tile groups and/or tiles.

Embodiments of the video decoder 30 as shown in Fig. 3 may be configured to partition
5 and/or decode the picture by using slices (also referred to as video slices), wherein a picture may be partitioned into or decoded using one or more slices (typically non-overlapping), and each slice may comprise one or more blocks (e.g. CTUs) or one or more groups of blocks (e.g. tiles (H.265/HEVC and VVC) or bricks (VVC)).

10 Embodiments of the video decoder 30 as shown in Fig. 3 may be configured to partition and/or decode the picture by using slices/tile groups (also referred to as video tile groups) and/or tiles (also referred to as video tiles), wherein a picture may be partitioned into or decoded using one or more slices/tile groups (typically non-overlapping), and each slice/tile group may comprise, e.g. one or more blocks (e.g. CTUs) or one or more tiles, wherein each
15 tile, e.g. may be of rectangular shape and may comprise one or more blocks (e.g. CTUs), e.g. complete or fractional blocks.

Other variations of the video decoder 30 can be used to decode the encoded picture data 21. For example, the decoder 30 can produce the output video stream without the loop filtering
20 unit 320. For example, a non-transform based decoder 30 can inverse-quantize the residual signal directly without the inverse-transform processing unit 312 for certain blocks or frames. In another implementation, the video decoder 30 can have the inverse-quantization unit 310 and the inverse-transform processing unit 312 combined into a single unit.

25 It should be understood that, in the encoder 20 and the decoder 30, a processing result of a current step may be further processed and then output to the next step. For example, after interpolation filtering, motion vector derivation or loop filtering, a further operation, such as Clip or shift, may be performed on the processing result of the interpolation filtering, motion vector derivation or loop filtering.

30

It should be noted that further operations may be applied to the derived motion vectors of current block (including but not limit to control point motion vectors of affine mode, sub-block motion vectors in affine, planar, ATMVP modes, temporal motion vectors, and so on). For example, the value of motion vector is constrained to a predefined range according to its
35 representing bit. If the representing bit of motion vector is bitDepth, then the range is -

2^{^(bitDepth-1)} ~ 2^{^(bitDepth-1)}-1, where “^” means exponentiation. For example, if bitDepth is set equal to 16, the range is -32768 ~ 32767; if bitDepth is set equal to 18, the range is -131072~131071. For example, the value of the derived motion vector (e.g. the MVs of four 4x4 sub-blocks within one 8x8 block) is constrained such that the max difference between integer parts of the four 4x4 sub-block MVs is no more than N pixels, such as no more than 1 pixel. Here provides two methods for constraining the motion vector according to the bitDepth.

FIG. 4 is a schematic diagram of a video coding device 400 according to an embodiment of the disclosure. The video coding device 400 is suitable for implementing the disclosed
 10 embodiments as described herein. In an embodiment, the video coding device 400 may be a decoder such as video decoder 30 of FIG. 1A or an encoder such as video encoder 20 of FIG. 1A.

The video coding device 400 comprises ingress ports 410 (or input ports 410) and receiver units (Rx) 420 for receiving data; a processor, logic unit, or central processing unit (CPU) 430 to process the data; transmitter units (Tx) 440 and egress ports 450 (or output ports 450) for transmitting the data; and a memory 460 for storing the data. The video coding device 400 may also comprise optical-to-electrical (OE) components and electrical-to-optical (EO) components coupled to the ingress ports 410, the receiver units 420, the transmitter units 440,
 20 and the egress ports 450 for egress or ingress of optical or electrical signals.

The processor 430 is implemented by hardware and software. The processor 430 may be implemented as one or more CPU chips, cores (e.g., as a multi-core processor), FPGAs, ASICs, and DSPs. The processor 430 is in communication with the ingress ports 410,
 25 receiver units 420, transmitter units 440, egress ports 450, and memory 460. The processor 430 comprises a coding module 470. The coding module 470 implements the disclosed embodiments described above. For instance, the coding module 470 implements, processes, prepares, or provides the various coding operations. The inclusion of the coding module 470 therefore provides a substantial improvement to the functionality of the video coding
 30 device 400 and effects a transformation of the video coding device 400 to a different state. Alternatively, the coding module 470 is implemented as instructions stored in the memory 460 and executed by the processor 430.

The memory 460 may comprise one or more disks, tape drives, and solid-state drives and
 35 may be used as an over-flow data storage device, to store programs when such programs are selected for execution, and to store instructions and data that are read during program

execution. The memory 460 may be, for example, volatile and/or non-volatile and may be a read-only memory (ROM), random access memory (RAM), ternary content-addressable memory (TCAM), and/or static random-access memory (SRAM).

- 5 Fig. 5 is a simplified block diagram of an apparatus **500** that may be used as either or both of the source device 12 and the destination device 14 from Fig. 1 according to an exemplary embodiment.

A processor 502 in the apparatus 500 can be a central processing unit. Alternatively, the processor 502 can be any other type of device, or multiple devices, capable of manipulating
10 or processing information now-existing or hereafter developed. Although the disclosed implementations can be practiced with a single processor as shown, e.g., the processor 502, advantages in speed and efficiency can be achieved using more than one processor.

A memory 504 in the apparatus 500 can be a read only memory (ROM) device or a random
15 access memory (RAM) device in an implementation. Any other suitable type of storage device can be used as the memory 504. The memory 504 can include code and data 506 that is accessed by the processor 502 using a bus 512. The memory 504 can further include an operating system 508 and application programs 510, the application programs 510 including at least one program that permits the processor 502 to perform the methods described here.
20 For example, the application programs 510 can include applications 1 through N, which further include a video coding application that performs the methods described here. The apparatus 500 can also include one or more output devices, such as a display 518. The display 518 may be, in one example, a touch sensitive display that combines a display with a touch sensitive element that is operable to sense touch inputs. The display 518 can be coupled
25 to the processor 502 via the bus 512.

Although depicted here as a single bus, the bus 512 of the apparatus 500 can be composed of multiple buses. Further, the secondary storage 514 can be directly coupled to the other
30 components of the apparatus 500 or can be accessed via a network and can comprise a single integrated unit such as a memory card or multiple units such as multiple memory cards. The apparatus 500 can thus be implemented in a wide variety of configurations.

Embodiments of the current invention allow to implement two or more reference picture management schemes (e.g. implicit sliding window method (e.g. as used in AVC) and reference picture list (RPL) management method (e.g. VVC) inside one codec.

5

In a first embodiment of the current invention, implicit sliding window method and Reference Picture Lists (RPL) methods are combined into one codec and any of these two methods can be selected, e.g., for usage at the sequence (SPS) level.

- 10 The following table provides a snippet of a SPS syntax table specifying the reference picture management method switch.

seq_parameter_set_rbsp() {	Descriptor
...	
sps_rpl_flag	u(1)
...	
}	

sps_rpl_flag equal to 1 specifies that syntax related to reference picture lists is present and RPL method is used; sps_rpl_flag equal to 0 specifies that syntax related to reference picture

- 15 lists is not present and AVC method is used.

The following table provides an example of signalling reference pictures syntax related to AVC or RPL methods.

seq_parameter_set_rbsp() {	Descriptor
...	
if(!sps_rpl_flag)	
max_num_tid0_ref_pics	ue(v)
else {	
long_term_ref_pics_flag	u(1)
rpl1_same_as_rpl0_flag	u(1)
for(i = 0; i < !rpl1_same_as_rpl0_flag ? 2 : 1; i++) {	
num_ref_pic_lists_in_sps[i]	ue(v)
for(j = 0; j < num_ref_pic_lists_in_sps[i]; j++)	
ref_pic_list_struct(i, j, long_term_ref_pics_flag)	
}	
...	
}	
picture_cropping_flag	u(1)
if(picture_cropping_flag) {	

The syntax parameter (or flag) `sps_rpl_flag` is, e.g., sent at SPS level (in further embodiment it might be sent at PPS or slice header) and indicates whether RPL shall be used for reference picture management (flag value “1” indicates RPL to be used, flag value “0” indicates that RPL is not to be used but the implicit sliding window mechanism).

If RPL method is not enabled (`sps_rpl_flag` equal to 0), only `max_num_tid0_ref_pics` is signalled for the implicit sliding window reference picture management method.

`max_num_tid0_ref_pics` specifies the maximum number of reference pictures in temporal layer 0 that may be used by the decoding process for inter prediction (implicit sliding window reference picture management method) of any picture in the CVS (Coding Video sequence).

The value of `max_num_tid0_ref_pics` is also used to determine the size of the Decoded Picture Buffer. The value of `max_num_tid0_ref_pics` is typically in the range of 0 to 5. When

not present, the value of `max_num_tid0_ref_pics` is inferred, e.g., to be equal to 0.

`long_term_ref_pics_flag` equal to 0 specifies that no LTRP is used for inter prediction of any coded picture in the CVS. `long_term_ref_pics_flag` equal to 1 specifies that LTRPs may be used for inter prediction of one or more coded pictures in the CVS. When not present, the value of `long_term_ref_pics_flag` is inferred to be equal to 0.

`rpl1_same_as_rpl0_flag` equal to 1 specifies that the syntax structures `num_ref_pic_lists_in_sps[ 1 ]` and `ref_pic_list_struct( 1, rplIdx, ltrpFlag )` are not present and the following applies:

— The value of `num_ref_pic_lists_in_sps[ 1 ]` is inferred to be equal to the value of `num_ref_pic_lists_in_sps[ 0 ]`.

— The value of each of syntax elements in `ref_pic_list_struct( 1, rplIdx, ltrpFlag )` is inferred to be equal to the value of corresponding syntax element in `ref_pic_list_struct( 0, rplIdx, ltrpFlag )` for `rplIdx` ranging from 0 to `num_ref_pic_lists_in_sps[ 0 ] - 1`.

`num_ref_pic_lists_in_sps[ i ]` specifies the number of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structures with `listIdx` equal to `i` included in the SPS. The value of `num_ref_pic_lists_in_sps[ i ]` shall be in the range of 0 to 64, inclusive. When not present, the value of `num_ref_pic_lists_in_sps[ i ]` is inferred to be equal to 0.

NOTE For each value of `listIdx` (equal to 0 or 1), a decoder must allocate memory for a total number of `num_ref_pic_lists_in_sps[ i ] + 1` `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structures since there can be one `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure directly signalled in the slice headers of a current picture.

If RPL method is enabled (`sps_rpl_flag` equal to 1), the following RPL syntax is signalled at the slice level.

<code>slice_header() {</code>	Descriptor
<code>...</code>	
<code>if(sps_rpl_flag) {</code>	
<code> for(i = 0; i < 2; i++) {</code>	

if(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

According to the slice header syntax table, there is a syntax element (e.g.

ref_pic_list_sps_flag[i]) specifying whether reference picture list is signalled in the slice header or some of reference picture lists that are already signalled at SPS level, are used in the

5 current slice.

In the current embodiment **ref_pic_list_sps_flag[i]** is always signalled in the slice header for list0 and also signalled for list1 only if rpl1_idx_present_flag is equal to 1.

- 10 In another embodiment similar to the first embodiment, if RPL method is enabled (sps_rpl_flag equal to 1), the following RPL syntax is signalled at the slice level.

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag[i]	u(1)
if(ref_pic_list_sps_flag[i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx[i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

In this embodiment signalling of `ref_pic_list_sps_flag[ i ]` element is not performed if number of RPL lists inside SPS is equal 0.

`ref_pic_list_sps_flag[ i ]` equal to 1 specifies that reference picture list `i` of the current picture is derived based on one of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structures with `listIdx` equal to `i` in the active SPS. `ref_pic_list_sps_flag[ i ]` equal to 0 specifies that reference picture list `i` of the current picture is derived based on the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure with `listIdx` equal to `i` that is directly included in the slice headers of the current picture. When `num_ref_pic_lists_in_sps[ i ]` is equal to 0, the value of `ref_pic_list_sps_flag[ i ]` shall be equal to 0. When `rpl1_idx_present_flag` is equal to 0 and `ref_pic_list_sps_flag[ 0 ]` is present, the value of `ref_pic_list_sps_flag[ 1 ]` is inferred to be equal to `ref_pic_list_sps_flag[ 0 ]`. When not present, the value of `ref_pic_list_sps_flag[ i ]` is inferred to be equal to 0.

`ref_pic_list_idx[ i ]` specifies the index, into the list of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structures with `listIdx` equal to `i` included in the active SPS, of the `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag )` syntax structure with `listIdx` equal to `i` that is used for derivation of reference picture list `i` of the current picture. The syntax element `ref_pic_list_idx[ i ]` is represented by $\text{Ceil}(\text{Log2}(\text{num_ref_pic_lists_in_sps}[i]))$ bits. When not present, the value of `ref_pic_list_idx[ i ]` is inferred to be equal to 0. The value of `ref_pic_list_idx[ i ]` shall be in the range of 0 to `num_ref_pic_lists_in_sps[ i ] - 1`, inclusive. When `rpl1_idx_present_flag` is equal to 0 and `ref_pic_list_sps_flag[ 0 ]` is present, the value of `ref_pic_list_idx[ 1 ]` is inferred to be equal to `ref_pic_list_idx[ 0 ]`.

Following is an explanation of the applications of the encoding method as well as the decoding method as shown in the above-mentioned embodiments, and a system using them.

FIG. 6 is a block diagram showing a content supply system 3100 for realizing content distribution service. This content supply system 3100 includes capture device 3102, terminal device 3106, and optionally includes display 3126. The capture device 3102 communicates with the terminal device 3106 over communication link 3104. The communication link may include the communication channel 13 described above. The communication link 3104 includes but not limited to WIFI, Ethernet, Cable, wireless (3G/4G/5G), USB, or any kind of combination thereof, or the like.

The capture device 3102 generates data, and may encode the data by the encoding method as shown in the above embodiments. Alternatively, the capture device 3102 may distribute the data to a streaming server (not shown in the Figures), and the server encodes the data and transmits the encoded data to the terminal device 3106. The capture device 3102 includes but not limited to camera, smart phone or Pad, computer or laptop, video conference system, PDA, vehicle mounted device, or a combination of any of them, or the like. For example, the capture device 3102 may include the source device 12 as described above. When the data includes video, the video encoder 20 included in the capture device 3102 may actually perform video encoding processing. When the data includes audio (i.e., voice), an audio encoder included in the capture device 3102 may actually perform audio encoding processing. For some practical

scenarios, the capture device 3102 distributes the encoded video and audio data by multiplexing them together. For other practical scenarios, for example in the video conference system, the encoded audio data and the encoded video data are not multiplexed. Capture device 3102 distributes the encoded audio data and the encoded video data to the terminal device 3106 separately.

In the content supply system 3100, the terminal device 310 receives and reproduces the encoded data. The terminal device 3106 could be a device with data receiving and recovering capability, such as smart phone or Pad 3108, computer or laptop 3110, network video recorder (NVR)/ digital video recorder (DVR) 3112, TV 3114, set top box (STB) 3116, video conference system 3118, video surveillance system 3120, personal digital assistant (PDA) 3122, vehicle mounted device 3124, or a combination of any of them, or the like capable of decoding the above-mentioned encoded data. For example, the terminal device 3106 may include the destination device 14 as described above. When the encoded data includes video, the video decoder 30 included in the terminal device is prioritized to perform video decoding. When the encoded data includes audio, an audio decoder included in the terminal device is prioritized to perform audio decoding processing.

For a terminal device with its display, for example, smart phone or Pad 3108, computer or laptop 3110, network video recorder (NVR)/ digital video recorder (DVR) 3112, TV 3114, personal digital assistant (PDA) 3122, or vehicle mounted device 3124, the terminal device can feed the decoded data to its display. For a terminal device equipped with no display, for example, STB 3116, video conference system 3118, or video surveillance system 3120, an external display 3126 is contacted therein to receive and show the decoded data.

When each device in this system performs encoding or decoding, the picture encoding device or the picture decoding device, as shown in the above-mentioned embodiments, can be used.

FIG. 7 is a diagram showing a structure of an example of the terminal device 3106. After the terminal device 3106 receives stream from the capture device 3102, the protocol proceeding unit 3202 analyzes the transmission protocol of the stream. The protocol includes but not limited to Real Time Streaming Protocol (RTSP), Hyper Text Transfer Protocol (HTTP), HTTP Live streaming protocol (HLS), MPEG-DASH, Real-time Transport protocol (RTP), Real Time Messaging Protocol (RTMP), or any kind of combination thereof, or the like.

After the protocol proceeding unit 3202 processes the stream, stream file is generated. The file is outputted to a demultiplexing unit 3204. The demultiplexing unit 3204 can separate the multiplexed data into the encoded audio data and the encoded video data. As described above, for some practical scenarios, for example in the video conference system, the encoded audio

data and the encoded video data are not multiplexed. In this situation, the encoded data is transmitted to video decoder 3206 and audio decoder 3208 without through the demultiplexing unit 3204.

- 5 Via the demultiplexing processing, video elementary stream (ES), audio ES, and optionally subtitle are generated. The video decoder 3206, which includes the video decoder 30 as explained in the above mentioned embodiments, decodes the video ES by the decoding method as shown in the above-mentioned embodiments to generate video frame, and feeds this data to the synchronous unit 3212. The audio decoder 3208, decodes the audio ES to generate audio
10 frame, and feeds this data to the synchronous unit 3212. Alternatively, the video frame may store in a buffer (not shown in FIG. 7) before feeding it to the synchronous unit 3212. Similarly, the audio frame may store in a buffer (not shown in FIG. 7) before feeding it to the synchronous unit 3212.

The synchronous unit 3212 synchronizes the video frame and the audio frame, and supplies the
15 video/audio to a video/audio display 3214. For example, the synchronous unit 3212 synchronizes the presentation of the video and audio information. Information may code in the syntax using time stamps concerning the presentation of coded audio and visual data and time stamps concerning the delivery of the data stream itself.

If subtitle is included in the stream, the subtitle decoder 3210 decodes the subtitle, and
20 synchronizes it with the video frame and the audio frame, and supplies the video/audio/subtitle to a video/audio/subtitle display 3216.

In accordance with the description above, in particular, the following embodiments are provided. FIG. 8 illustrates a method of decoding an encoded video sequence according to an
25 embodiment. The method illustrated in Fig. 8 comprises the step of obtaining 810 (for example, by parsing) a value of a reference picture list (RPL) syntax element (e.g. sps_rpl_flag) from a bitstream of the video sequence. The RPL syntax element specifies whether syntax elements related to reference picture lists are present in the bitstream. The method, further, comprises, when the RPL syntax element specifies that the syntax elements related to reference picture
30 lists are present in the bitstream, obtaining 820, from the bitstream, values of the syntax elements related to reference picture lists. At least one reference picture list for inter prediction using the obtained values of the syntax elements related to reference picture lists is constructed 830, when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream.

The method may further comprise, when the RPL syntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, constructing at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream. In this case, construction of at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream may comprise performing a sliding window method to construct at least one reference picture list for inter prediction.

Thus, reference picture list management that is based and that is not based on reference picture list can be performed. By means of the RPL syntax element (for example, an “sps_rpl_flag”), whether a reference picture lists method can be used for an inter prediction process or not, and, thus, the management of the inter prediction process, particularly, the reference picture management, can be efficiently controlled.

The steps of constructing the at least one reference picture list for inter prediction using the obtained values of the syntax elements related to reference picture lists and at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream, respectively, can be performed in accordance with the EVC standard known to the skilled person. With reference to that standard, when at least one reference picture list for inter prediction using the obtained values of the syntax elements related to reference picture lists is constructed the following applies for the construction process:

“Reference pictures are addressed through reference indices. A reference index is an index into a reference picture list. When decoding an I slice, no reference picture list is used in decoding of the slice data. When decoding a P slice, only reference picture list 0 (i.e., RefPicList[0]) is used in decoding of the slice data. When decoding a B slice, both reference picture list 0 (i.e., RefPicList[0]) and the reference picture list 1 (i.e., RefPicList[1]) are used in decoding of the slice data.

At the beginning of the decoding process for each slice of a non-IDR picture, the reference picture lists RefPicList[0] and RefPicList[1] are derived. The reference picture lists are used in marking of reference pictures as specified in subclause 8.3.3 or in decoding of the slice data.

NOTE For an I slice of a non-IDR picture that it is not the first slice of the picture, RefPicList[0] and RefPicList[1] can be derived for bitstream conformance checking purpose, but their derivation is not necessary for decoding of the current picture or pictures following the current picture in decoding order. For a P slice that it is not the first slice of a picture, RefPicList[1] can be derived for bitstream conformance checking purpose, but its derivation is not necessary for decoding of the current picture or pictures following the current picture in decoding order.

The reference picture lists RefPicList[0] and RefPicList[1] are constructed as follows:

```

    for( i = 0; i < 2; i++ )
        for( j = 0, pocBase = PicOrderCntVal; j < NumEntriesInList[ i ][ SliceRplsIdx[ i ] ]; j++ ) {
            if( !lt_ref_pic_flag[ i ][ SliceRplsIdx[ i ] ][ j ] ) {
                RefPicPocList[ i ][ j ] = pocBase - DeltaPocSt[ i ][ SliceRplsIdx[ i ] ][ j ]
                if( there is a reference picture picA in the DPB with PicOrderCntVal equal to
                    RefPicPocList[ i ][ j ] )
                    RefPicList[ i ][ j ] = picA
                else
                    RefPicList[ i ][ j ] = "no reference picture"
                pocBase = RefPicPocList[ i ][ j ]
            } else {
                if( there is a reference picA in the DPB with PicOrderCntVal & ( MaxLtPicOrderCntLsb - 1 )
                    equal to FullPocLsbLt[ i ][ SliceRplsIdx[ i ] ][ j ] )
                    RefPicList[ i ][ j ] = picA
                else
                    RefPicList[ i ][ j ] = "no reference picture"
            }
        }

```

- 20 For each i equal to 0 or 1, the following applies:
- The first NumRefIdxActive[i] entries in RefPicList[i] are referred to as the active entries in RefPicList[i], and the other entries in RefPicList[i] are referred to as the inactive entries in RefPicList[i].
 - Each entry in RefPicList[i][j] for j in the range of 0 to NumEntriesInList[i][SliceRplsIdx[i]] - 1, inclusive, is referred to as an STRP entry if lt_ref_pic_flag[i][SliceRplsIdx[i]][j] is equal to 0, and as an LTRP entry otherwise."

With continued reference to that standard, when at least one reference picture list for inter prediction not using the obtained values of the syntax elements related to reference picture lists is constructed the following applies for the construction process:

- 30 "
- 1) The decoding process for filling a reference picture list with lower PicOrderCntVal in subclause 8.3.2.2.2 is invoked with i set equal to 0 and startIdx set equal to 0, and the output is the variable nextIdx.
 - 2) When nextIdx is less than NumRefIdxActive[0], the decoding process for filling a reference picture list with higher PicOrderCntVal in subclause 8.3.2.2.3 is invoked with i set equal to 0 and startIdx set equal to nextIdx, and the output is the variable nextIdx.
 - 3) When nextIdx is less than NumRefIdxActive[0], NumRefIdxActive[0] is set equal to nextIdx.

For B slices, the reference picture list RefPicList[1] is constructed as follows:

- 1) The decoding process for filling a reference picture list with higher PicOrderCntVal in subclause 8.3.2.2.3 is invoked with i set equal to 1 and startIdx set equal to 0, and the output is the variable nextIdx.
- 2) When nextIdx is less than NumRefIdxActive[1], the decoding process for filling a reference picture list with lower PicOrderCntVal in subclause 8.3.2.2.2 is invoked with i set equal to 1 and startIdx set equal to nextIdx, and the output is the variable nextIdx.
- 3) When nextIdx is less than NumRefIdxActive[1], NumRefIdxActive[1] is set equal to nextIdx."

8.3.2.2.2 Decoding process for filling a reference picture list with lower PicOrderCntVal pictures

Inputs to this process are:

— a reference picture list identifier i , and

5 — a start index position $startIdx$.

Output of this process is the variable $nextIdx$ representing the number of positions filled in the reference picture list.

The variable $nextIdx$ is set equal to $startIdx$.

The variable $nextTemporalId$ is set equal to $\text{Max}(\text{TemporalId} - 1, 0)$.

10 Let $minPoc$ be set equal to the lowest value of $\text{PictureOrderCountVal}$ of all reference pictures in the DPB.

The reference picture list $\text{RefPicList}[i]$ is filled with lower PicOrderCntVal pictures as follows:

```

    for( j = PicOrderCntVal; j >= minPoc && nextIdx < NumRefIdxActive[ i ]; j-- ) {
        if( there is a reference picture picA in the DPB with PicOrderCntVal equal to j and
            with TemporalId <= nextTemporalId ) {
15             RefPicList[ i ][ nextIdx++ ] = picA
                nextTemporalId = Max( TemporalId of picA - 1, 0 )
            }
        }
    }

```

(167)

8.3.2.2.3 Decoding process for filling a reference picture list with higher PicOrderCntVal

20 Inputs to this process are:

— a reference picture list identifier i , and

— a start index position $startIdx$.

Output of this process is the variable $nextIdx$ representing the number of positions filled in the reference picture list.

25 The variable $nextIdx$ is set equal to $startIdx$.

The variable $nextTemporalId$ is set equal to $\text{Max}(\text{TemporalId} - 1, 0)$.

Let $maxPoc$ be set equal to the highest value of $\text{PictureOrderCountVal}$ of all reference pictures in the DPB.

The reference picture list $\text{RefPicList}[i]$ is filled with higher PicOrderCntVal pictures as follows:

```

    for( j = PicOrderCntVal; j <= maxPoc && nextIdx < NumRefIdxActive[ i ]; j++ ) {
30         if( there is a reference picture picA in the DPB with PicOrderCntVal equal to j and
            with TemporalId <= nextTemporalId ) {
                RefPicList[ i ][ nextIdx++ ] = picA
                nextTemporalId = Max( TemporalId of picA - 1, 0 )
            }
        }
35     }

```

(168)

According to another embodiment, it is provided a method of encoding a video sequence corresponding to the above-described decoding method and as illustrated in Fig. 9. The method of encoding a video sequence illustrated in Fig. 9 comprises the step of determining 910 whether values of syntax elements related to reference picture lists are to be used for construction (by a decoding device) of at least one reference picture list for inter prediction. Further, this method comprises generating 920 a bitstream comprising a RPL syntax element

specifying whether syntax elements related to reference picture lists are present in the bitstream, wherein the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream in case that it is determined that the values of the syntax elements related to reference picture lists are to be used (by the decoding device) for construction of at least one reference picture list for inter prediction.

The above-described methods may be implemented in a decoding device or encoding device, respectively, as described in the following.

As shown in Fig. 10 it is provided, according to an embodiment, a decoding device 1000 for decoding an encoded video sequence. The decoding device comprises an obtaining unit 1010 configured for obtaining a value of a reference picture list (RPL) syntax element from a bitstream of the video sequence, the RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream and, when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream, obtaining, from the bitstream, values of the syntax elements related to reference picture lists. Further, the decoding device 1000 comprises a RPL construction unit 1020 for, when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream, constructing at least one reference picture list for inter prediction using the values of the syntax elements related to reference picture lists provided by the obtaining unit 1010.

Similarly, it is provided an encoding device for encoding a video sequence as shown in Fig. 11. The encoding device shown in Fig. 11 comprises a determining unit 1110 for determining whether values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction. Further, the encoding device 1100 comprises a bitstream generation unit 1120 configured for generating a bitstream comprising a RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream, wherein the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream in case that it is determined by the determining unit 1110 that the values of the syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

The decoding device 1000 shown in Fig. 10 may be or may be comprised by the decoder 30 shown in Fig. 1A, 1B, 3 and the video decoder 3206 shown in Fig. 7. Further, the decoding device 1000 may be comprised by the video coding device 400 shown in Fig. 4, the apparatus 500 shown in Fig. 5 and the terminal device 3106 shown in Fig. 6. The encoding device 1100 shown in Fig. 11 may be or may be comprised by the encoder 20 shown in Fig. 1A, 1B and 3. Further, the encoding device 1100 may be comprised by the video coding device 400 shown in Fig. 4, the apparatus 500 shown in Fig. 5 and the capture device 3102 shown in Fig. 6.

Furthermore, the following embodiments are provided herein.

1. A method for decoding (e.g. a video), e.g. implemented by a decoding device, comprising: obtaining a flag (e.g. `sps_rpl_flag` or some other parameter) specifying whether syntax (e.g. syntax information or syntax parameters) related to reference picture lists (e.g. RPL) is present in a bitstream by parsing the bitstream (e.g. of an encoded video);
 - if the flag (e.g. `sps_rpl_flag`) specifying whether syntax (e.g. syntax information or syntax parameters) related to reference picture lists (e.g. RPL) is present in the bitstream, is equal to 1 (or in general, if the flag value is set to a first value, e.g. 1 or 0, indicating that RPL shall be used, wherein a second opposite/inverse value, e.g. 0 or 1, indicates that RPL shall not be used)
 - obtaining the syntax elements representing (or defining) the reference picture lists by parsing the bitstream; and
 - performing reference picture list management based on the obtained reference picture lists;
 - otherwise,
 - performing a different reference picture list management (e.g. a reference picture management using an implicit sliding window method, e.g. using syntax element `max_num_tid0_ref_pics`).
2. The method of embodiment 1, wherein the flag (e.g. `sps_rpl_flag`) specifying whether syntax related to reference picture lists (e.g. RPL) is present in a bitstream is signaled at a SPS level.
3. The method of embodiment 1 or 2, wherein the syntax (syntax information or syntax parameters) representing the reference picture lists is signalled at a sequence parameter set level (SPS) (at SPS level e.g. `rpl1_same_as_rpl0_flag` and `num_ref_pic_lists_in_sps[i]`, where “i” is a number of reference picture list, e.g. value “0” corresponds to list0 and value “1”

corresponds to list1) and/or a slice header (SH) level (at SH level e.g. `ref_pic_list_sps_flag[ i ]` and `ref_pic_list_idx[ i ]`, where “i” is a number of reference picture list, e.g. value “0” corresponds to list0 and value “1” corresponds to list1).

- 5 4. The method of any of the preceding embodiments, wherein a syntax element (e.g. **`ref_pic_list_sps_flag[ i ]`**), signalled at the slice header level (or in a slice header), specifies whether the reference picture lists (see e.g. syntax table `ref_pic_list_struct( listIdx, rplIdx, ltrpFlag)`) that are signalled at SPS level, are used in or for the current slice (to which the slice header belongs to).

10

5. The method of embodiment 4, wherein the syntax element (e.g. **`ref_pic_list_sps_flag[ i ]`**), signalled at slice header level, specifying whether the reference picture lists that are signalled at SPS level are used in the current slice, is always signalled for a first reference list (e.g. list0) and also signalled for a second reference list (e.g. list1) only if `rpl1_idx_present_flag` is equal to 1 (or in general only if a syntax element, e.g. a second flag, indicating whether a second reference list is signaled in the bitstream has a predetermined value, e.g. 1).
- 15

6. The method of embodiment 5, wherein the syntax element (e.g. **`ref_pic_list_sps_flag[ i ]`**), signalled at the slice header level (or in a slice header), specifying whether the reference picture lists that are signalled at SPS level, are used in the current slice (to which the slice header belongs to), is signalled at the slice header according to the following table snippet.
- 20

slice_header() {	Descriptor
...	
if(<code>sps_rpl_flag</code>) {	
for(<code>i = 0; i < 2; i++</code>) {	
if(<code>i == 0 (i == 1 && rpl1_idx_present_flag)</code>)	
<code>ref_pic_list_sps_flag[i]</code>	u(1)
if(<code>ref_pic_list_sps_flag[i]</code>) {	
if(<code>num_ref_pic_lists_in_sps[i] > 1 &&</code>	
<code>(i == 0 (i == 1 && rpl1_idx_present_flag))</code>)	
<code>ref_pic_list_idx[i]</code>	u(v)
} else	
<code>ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)</code>	
...	
}	

...	
}	

7. The method of embodiment 4, wherein the syntax element (e.g. **ref_pic_list_sps_flag**[i]),
signalled at a slice header level (or in a slice header), which specifies whether the reference
picture lists that are signalled at SPS level are used in or for the current slice, is not signalled
5 if there are no reference picture lists signalled at SPS level.

8. The method of embodiment 7, wherein the syntax element (e.g. **ref_pic_list_sps_flag**[i]),
signalled at the slice header level, which specifies whether the reference picture lists that are
signalled at SPS level are used in the current slice, is signalled for list0 and signalled for list1
10 only if **rpl1_idx_present_flag** is equal to 1 (or in general only if a syntax element, .e.g. a second
flag, indicating whether a second reference list is signaled in the bitstream has a predetermined
value, e.g. 1).

9. The method of embodiment 7 or embodiment 8, wherein the syntax element specifying
15 which reference picture list is used is signalled at the slice header according to the following
table snippet.

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_sps_flag [i]	u(1)
if(ref_pic_list_sps_flag [i]) {	
if(num_ref_pic_lists_in_sps[i] > 1 && (i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx [i]	u(v)
} else	
ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)	
...	
}	
...	
}	

10. A method for encoding (e.g. a video), e.g. implemented by an encoding device, comprising:
generating a bitstream (e.g. representing an encoded version of the video);
adding (to the bitstream) a flag (e.g. `sps_rpl_flag` or some other parameter) specifying
whether syntax (e.g. syntax information or syntax parameters) related to reference picture lists
5 (e.g. RPL) is present (or is added) in a bitstream;
- if the flag (e.g. `sps_rpl_flag`) specifying whether syntax (e.g. syntax information or
syntax parameters) related to reference picture lists (e.g. RPL) is present in the
bitstream, is equal to 1 (or in general, if the flag value is set to a first value, e.g. 1 or 0,
10 indicating that RPL shall be used, wherein a second opposite/inverse value, e.g. 0 or
1, indicates that RPL shall not be used)
 - adding the syntax elements representing (or defining) the reference picture lists to
the bitstream;
 - otherwise,
 - adding no or other syntax to enable the decoder to perform a different reference
15 picture list management (e.g. a reference picture management using an implicit
sliding window method, e.g. using syntax element `max_num_tid0_ref_pics`).
11. The method of embodiment 10, wherein the flag (e.g. `sps_rpl_flag`) specifying whether
20 syntax related to reference picture lists (e.g. RPL) is present in a bitstream is signaled at a SPS
level.
12. The method of embodiment 10 or 11, wherein the syntax (syntax information or syntax
parameters) representing the reference picture lists is signalled at a sequence parameter set
25 level (SPS) (e.g. at SPS level `rpl1_same_as_rpl0_flag` and `num_ref_pic_lists_in_sps[ i ]` ,
where “i” is a number of reference picture list, e.g. value “0” corresponds to list0 and value “1”
corresponds to list1) and/or a slice header (SH) level (at SH level e.g. `ref_pic_list_sps_flag[ i ]`
and `ref_pic_list_idx[ i ]` , where “i” is a number of reference picture list, e.g. value “0”
corresponds to list0 and value “1” corresponds to list1).
- 30
13. The method of any of the embodiments 10 to 12, wherein a syntax element (e.g.
`ref_pic_list_sps_flag[ i ]`), signalled at the slice header level (or in a slice header), specifies
whether the reference picture lists (see e.g. syntax table `ref_pic_list_struct( listIdx, rplIdx,`
`ltpFlag`)) that are signalled at SPS level, are used in or for the current slice (to which the slice
35 header belongs to).

14. The method of embodiment 13, wherein the syntax element (e.g. **ref_pic_list_sps_flag**[i]), signalled at slice header level, specifying whether the reference picture lists that are signalled at SPS level are used in the current slice, is always signalled for a first reference list (e.g. list0) and also signalled for a second reference list (e.g. list1) only if **rpl1_idx_present_flag** is equal to 1 (or in general only if a syntax element, e.g. a second flag, indicating whether a second reference list is signaled in the bitstream has a predetermined value, e.g. 1).

15. The method of embodiment 14, wherein the syntax element (e.g. **ref_pic_list_sps_flag**[i]), signalled at the slice header level (or in a slice header), specifying whether the reference picture lists that are signalled at SPS level are used in the current slice (to which the slice header belongs to), is signalled at the slice header according to the following table snippet.

slice_header() {	Descriptor
...	
if(sps_rpl_flag) {	
for(i = 0; i < 2; i++) {	
if(i == 0 (i == 1 && rpl1_idx_present_flag))	
ref_pic_list_sps_flag [i]	u(1)
if(ref_pic_list_sps_flag [i]) {	
if(num_ref_pic_lists_in_sps [i] > 1 &&	
(i == 0 (i == 1 && rpl1_idx_present_flag)))	
ref_pic_list_idx [i]	u(v)
} else	
ref_pic_list_struct (i, num_ref_pic_lists_in_sps [i], long_term_ref_pics_flag)	
...	
}	
...	
}	

16. The method of embodiment 15, wherein the syntax element (e.g. **ref_pic_list_sps_flag**[i]), signalled at a slice header level (or in a slice header), which specifies whether the reference picture lists that are signalled at SPS level, are used in or for the current slice, is not signalled if there are no reference picture lists signalled at SPS level .
17. The method of embodiment 16, wherein the syntax element (e.g. **ref_pic_list_sps_flag**[i]), signalled at the slice header level, which specifies whether the reference picture lists that are signalled at SPS level are used in the current slice, is signalled for list0 and signalled for list1

only if `rpl1_idx_present_flag` is equal to 1 (or in general only if a syntax element, e.g. a second flag, indicating whether a second reference list is signaled in the bitstream has a predetermined value, e.g. 1).

- 5 18. The method of embodiment 16 or embodiment 17, wherein the syntax element specifying which reference picture list is used is signalled at the slice header according to the following table snippet.

slice_header() {	Descriptor
...	
if(<code>sps_rpl_flag</code>) {	
for(<code>i = 0; i < 2; i++</code>) {	
if(<code>num_ref_pic_lists_in_sps[i] > 0 && (i == 0 (i == 1 && rpl1_idx_present_flag))</code>)	
<code>ref_pic_list_sps_flag[i]</code>	u(1)
if(<code>ref_pic_list_sps_flag[i]</code>) {	
if(<code>num_ref_pic_lists_in_sps[i] > 1 &&</code>	
<code>(i == 0 (i == 1 && rpl1_idx_present_flag))</code>)	
<code>ref_pic_list_idx[i]</code>	u(v)
} else	
<code>ref_pic_list_struct(i, num_ref_pic_lists_in_sps[i], long_term_ref_pics_flag)</code>	
...	
}	
...	
}	

10

19. An encoder (20) comprising processing circuitry for carrying out the method according to any of embodiments 10 to 18.

20. A decoder (30) comprising processing circuitry for carrying out the method according to any of embodiments 1 to 9.
- 15

21. A computer program product comprising program code for performing the method according to any one of the embodiments 1 to 18 when executed on a computer or a processor

20

22. A decoder, comprising:
one or more processors; and
a non-transitory computer-readable storage medium coupled to the processors and storing
programming for execution by the processors, wherein the programming, when executed by
5 the processors, configures the decoder to carry out the method according to any one of the
embodiments 1 to 9.

23. An encoder, comprising:
one or more processors; and
10 a non-transitory computer-readable storage medium coupled to the processors and storing
programming for execution by the processors, wherein the programming, when executed by
the processors, configures the encoder to carry out the method according to any one of the
embodiments 10 to 18.

15 24. A non-transitory computer-readable medium carrying a program code which, when
executed by a computer device, causes the computer device to perform the method of any one
of the embodiments 1 to 18.

20 Furthermore, it is provided:

1. A method of coding implemented by a decoding device, comprising:
obtaining a flag specifying whether syntax related to reference picture lists is present in a
bitstream by parsing the bitstream;
25 In condition that syntax related to reference picture lists is present, obtaining the syntax
related to reference picture lists by parsing the bitstream;
Predicting pictures related to the flag, based on the syntax related to reference picture lists.

2. A method of coding implemented by an encoding device, comprising:
30 Determining whether syntax related to reference picture lists is present in a bitstream by
parsing the bitstream;
In condition that syntax related to reference picture lists is present, encoding the syntax
related to reference picture lists and a flag specifying the syntax related to reference picture
lists is present in the bitstream into the bitstream.

The present invention is not limited to the above-mentioned system, and either the picture encoding device or the picture decoding device in the above-mentioned embodiments can be incorporated into other system, for example, a car system.

Mathematical Operators

The mathematical operators used in this application are similar to those used in the C programming language. However, the results of integer division and arithmetic shift operations are defined more precisely, and additional operations are defined, such as exponentiation and real-valued division. Numbering and counting conventions generally begin from 0, e.g., "the first" is equivalent to the 0-th, "the second" is equivalent to the 1-th, etc.

15

Arithmetic operators

The following arithmetic operators are defined as follows:

+	Addition
−	Subtraction (as a two-argument operator) or negation (as a unary prefix operator)
*	Multiplication, including matrix multiplication
x^y	Exponentiation. Specifies x to the power of y. In other contexts, such notation is used for superscripting not intended for interpretation as exponentiation.
/	Integer division with truncation of the result toward zero. For example, $7 / 4$ and $-7 / -4$ are truncated to 1 and $-7 / 4$ and $7 / -4$ are truncated to -1 .
÷	Used to denote division in mathematical equations where no truncation or rounding is intended.
$\frac{x}{y}$	Used to denote division in mathematical equations where no truncation or rounding is intended.
$\sum_{i=x}^y f(i)$	The summation of $f(i)$ with i taking all integer values from x up to and including y.
$x \% y$	Modulus. Remainder of x divided by y, defined only for integers x and y with $x \geq 0$ and $y > 0$.

Logical operators

The following logical operators are defined as follows:

20 $x \&\& y$ Boolean logical "and" of x and y
 $x || y$ Boolean logical "or" of x and y
 $!$ Boolean logical "not"

$x ? y : z$ If x is TRUE or not equal to 0, evaluates to the value of y ; otherwise, evaluates to the value of z .

Relational operators

5 The following relational operators are defined as follows:

	$>$	Greater than
	$>=$	Greater than or equal to
	$<$	Less than
	$<=$	Less than or equal to
10	$=$	Equal to
	$!=$	Not equal to

When a relational operator is applied to a syntax element or variable that has been assigned the value "na" (not applicable), the value "na" is treated as a distinct value for the syntax element or variable. The value "na" is considered not to be equal to any other value.

Bit-wise operators

The following bit-wise operators are defined as follows:

20	$\&$	Bit-wise "and". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0.
25	$ $	Bit-wise "or". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0.
30	$\wedge$	Bit-wise "exclusive or". When operating on integer arguments, operates on a two's complement representation of the integer value. When operating on a binary argument that contains fewer bits than another argument, the shorter argument is extended by adding more significant bits equal to 0.
	$x \gg y$	Arithmetic right shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values of y . Bits shifted into the most significant bits (MSBs) as a result of the right shift have a value equal to the MSB of x prior to the shift operation.
35	$x \ll y$	Arithmetic left shift of a two's complement integer representation of x by y binary digits. This function is defined only for non-negative integer values of y . Bits shifted into the least significant bits (LSBs) as a result of the left shift have a value equal to 0.

40 Assignment operators

The following arithmetic operators are defined as follows:

	$=$	Assignment operator
	$++$	Increment, i.e., $x++$ is equivalent to $x = x + 1$; when used in an array index, evaluates to the value of the variable prior to the increment operation.

- Decrement, i.e., $x--$ is equivalent to $x = x - 1$; when used in an array index, evaluates to the value of the variable prior to the decrement operation.
- += Increment by amount specified, i.e., $x += 3$ is equivalent to $x = x + 3$, and $x += (-3)$ is equivalent to $x = x + (-3)$.
- 5 -= Decrement by amount specified, i.e., $x -= 3$ is equivalent to $x = x - 3$, and $x -= (-3)$ is equivalent to $x = x - (-3)$.

Range notation

The following notation is used to specify a range of values:

- 10 $x = y..z$ x takes on integer values starting from y to z , inclusive, with x , y , and z being integer numbers and z being greater than y .

Mathematical functions

The following mathematical functions are defined:

- 15
$$\text{Abs}(x) = \begin{cases} x & ; \quad x \geq 0 \\ -x & ; \quad x < 0 \end{cases}$$

$\text{Asin}(x)$ the trigonometric inverse sine function, operating on an argument x that is in the range of -1.0 to 1.0 , inclusive, with an output value in the range of $-\pi/2$ to $\pi/2$, inclusive, in units of radians

- 20 $\text{Atan}(x)$ the trigonometric inverse tangent function, operating on an argument x , with an output value in the range of $-\pi/2$ to $\pi/2$, inclusive, in units of radians

$$\text{Atan2}(y, x) = \begin{cases} \text{Atan}\left(\frac{y}{x}\right) & ; \quad x > 0 \\ \text{Atan}\left(\frac{y}{x}\right) + \pi & ; \quad x < 0 \ \&\& \ y \geq 0 \\ \text{Atan}\left(\frac{y}{x}\right) - \pi & ; \quad x < 0 \ \&\& \ y < 0 \\ +\frac{\pi}{2} & ; \quad x = 0 \ \&\& \ y \geq 0 \\ -\frac{\pi}{2} & ; \quad \text{otherwise} \end{cases}$$

$\text{Ceil}(x)$ the smallest integer greater than or equal to x .

$\text{Clip1}_Y(x) = \text{Clip3}(0, (1 \ll \text{BitDepth}_Y) - 1, x)$

$\text{Clip1}_C(x) = \text{Clip3}(0, (1 \ll \text{BitDepth}_C) - 1, x)$

- 25
$$\text{Clip3}(x, y, z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad \text{otherwise} \end{cases}$$

$\text{Cos}(x)$ the trigonometric cosine function operating on an argument x in units of radians.

$\text{Floor}(x)$ the largest integer less than or equal to x .

$$\text{GetCurrMsb}(a, b, c, d) = \begin{cases} c + d & ; \quad b - a \geq d/2 \\ c - d & ; \quad a - b > d/2 \\ c & ; \quad \text{otherwise} \end{cases}$$

- 30 $\text{Ln}(x)$ the natural logarithm of x (the base- e logarithm, where e is the natural logarithm base constant 2.718 281 828...).

$\text{Log2}(x)$ the base-2 logarithm of x .

$\text{Log10}(x)$ the base-10 logarithm of x .

$$\text{Min}(x, y) = \begin{cases} x & ; \quad x \leq y \\ y & ; \quad x > y \end{cases}$$

$$\text{Max}(x, y) = \begin{cases} x & ; \quad x \geq y \\ y & ; \quad x < y \end{cases}$$

5 $\text{Round}(x) = \text{Sign}(x) * \text{Floor}(\text{Abs}(x) + 0.5)$

$$\text{Sign}(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x = 0 \\ -1 & ; \quad x < 0 \end{cases}$$

$\text{Sin}(x)$ the trigonometric sine function operating on an argument x in units of radians

$$\text{Sqrt}(x) = \sqrt{x}$$

$$\text{Swap}(x, y) = (y, x)$$

10 $\text{Tan}(x)$ the trigonometric tangent function operating on an argument x in units of radians

Order of operation precedence

When an order of precedence in an expression is not indicated explicitly by use of parentheses, the following rules apply:

- 15 – Operations of a higher precedence are evaluated before any operation of a lower precedence.
- Operations of the same precedence are evaluated sequentially from left to right.

- 20 The table below specifies the precedence of operations from highest to lowest; a higher position in the table indicates a higher precedence.

For those operators that are also used in the C programming language, the order of precedence used in this Specification is the same as used in the C programming language.

- 25 Table: Operation precedence from highest (at top of table) to lowest (at bottom of table)

operations (with operands x, y, and z)
"x++", "x--"
"!x", "-x" (as a unary prefix operator)
x^y
"x * y", "x / y", "x ÷ y", " $\frac{x}{y}$ ", "x % y"
"x + y", "x - y" (as a two-argument operator), " $\sum_{i=x}^y f(i)$ "
"x << y", "x >> y"
"x < y", "x <= y", "x > y", "x >= y"
"x == y", "x != y"
"x & y"
"x y"
"x && y"
"x y"
"x ? y : z"
"x..y"
"x = y", "x += y", "x -= y"

Text description of logical operations

In the text, a statement of logical operations as would be described mathematically in the following form:

```

5      if( condition 0 )
        statement 0
      else if( condition 1 )
        statement 1
      ...
10     else /* informative remark on remaining condition */
        statement n

```

may be described in the following manner:

... as follows / ... the following applies:

```

15      - If condition 0, statement 0
      - Otherwise, if condition 1, statement 1
      - ...
      - Otherwise (informative remark on remaining condition), statement n

```

20 Each "If ... Otherwise, if ... Otherwise, ..." statement in the text is introduced with "... as follows" or "... the following applies" immediately followed by "If ... ". The last condition of the "If ... Otherwise, if ... Otherwise, ..." is always an "Otherwise, ...". Interleaved "If ...

Otherwise, if ... Otherwise, ..." statements can be identified by matching "... as follows" or "... the following applies" with the ending "Otherwise, ...".

5 In the text, a statement of logical operations as would be described mathematically in the following form:

```

    if( condition 0a  &&  condition 0b )
        statement 0
    else if( condition 1a  ||  condition 1b )
        statement 1
10  ...
    else
        statement n

```

may be described in the following manner:

... as follows / ... the following applies:

- 15 – If all of the following conditions are true, statement 0:
- condition 0a
 - condition 0b
- Otherwise, if one or more of the following conditions are true, statement 1:
- condition 1a
 - 20 – condition 1b
 - ...
 - Otherwise, statement n

In the text, a statement of logical operations as would be described mathematically in the following form:

```

25  if( condition 0 )
        statement 0
    if( condition 1 )
        statement 1

```

may be described in the following manner:

- 30 When condition 0, statement 0
- When condition 1, statement 1

Although embodiments of the invention have been primarily described based on video coding, it should be noted that embodiments of the coding system 10, encoder 20 and decoder 30 (and correspondingly the system 10) and the other embodiments described herein may also be configured for still picture processing or coding, i.e. the processing or coding of an individual picture independent of any preceding or consecutive picture as in video coding. In general only inter-prediction units 244 (encoder) and 344 (decoder) may not be available in case the picture processing coding is limited to a single picture 17. All other functionalities (also referred to as tools or technologies) of the video encoder 20 and video decoder 30 may equally be used for still picture processing, e.g. residual calculation 204/304, transform 206, quantization 208, inverse quantization 210/310, (inverse) transform 212/312, partitioning 262/362, intra-prediction 254/354, and/or loop filtering 220, 320, and entropy coding 270 and entropy decoding 304.

Embodiments, e.g. of the encoder 20 and the decoder 30, and functions described herein, e.g. with reference to the encoder 20 and the decoder 30, may be implemented in hardware, software, firmware, or any combination thereof. If implemented in software, the functions may be stored on a computer-readable medium or transmitted over communication media as one or more instructions or code and executed by a hardware-based processing unit. Computer-readable media may include computer-readable storage media, which corresponds to a tangible medium such as data storage media, or communication media including any medium that facilitates transfer of a computer program from one place to another, e.g., according to a communication protocol. In this manner, computer-readable media generally may correspond to (1) tangible computer-readable storage media which is non-transitory or (2) a communication medium such as a signal or carrier wave. Data storage media may be any available media that can be accessed by one or more computers or one or more processors to retrieve instructions, code and/or data structures for implementation of the techniques described in this disclosure. A computer program product may include a computer-readable medium.

By way of example, and not limiting, such computer-readable storage media can comprise RAM, ROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage, or other magnetic storage devices, flash memory, or any other medium that can be used to store desired program code in the form of instructions or data structures and that can be accessed by a computer. Also, any connection is properly termed a computer-readable medium. For

example, if instructions are transmitted from a website, server, or other remote source using a coaxial cable, fiber optic cable, twisted pair, digital subscriber line (DSL), or wireless technologies such as infrared, radio, and microwave, then the coaxial cable, fiber optic cable, twisted pair, DSL, or wireless technologies such as infrared, radio, and microwave are included in the definition of medium. It should be understood, however, that computer-readable storage media and data storage media do not include connections, carrier waves, signals, or other transitory media, but are instead directed to non-transitory, tangible storage media. Disk and disc, as used herein, includes compact disc (CD), laser disc, optical disc, digital versatile disc (DVD), floppy disk and Blu-ray disc, where disks usually reproduce data magnetically, while discs reproduce data optically with lasers. Combinations of the above should also be included within the scope of computer-readable media.

Instructions may be executed by one or more processors, such as one or more digital signal processors (DSPs), general purpose microprocessors, application specific integrated circuits (ASICs), field programmable logic arrays (FPGAs), or other equivalent integrated or discrete logic circuitry. Accordingly, the term “processor,” as used herein may refer to any of the foregoing structure or any other structure suitable for implementation of the techniques described herein. In addition, in some aspects, the functionality described herein may be provided within dedicated hardware and/or software modules configured for encoding and decoding, or incorporated in a combined codec. Also, the techniques could be fully implemented in one or more circuits or logic elements.

The techniques of this disclosure may be implemented in a wide variety of devices or apparatuses, including a wireless handset, an integrated circuit (IC) or a set of ICs (e.g., a chip set). Various components, modules, or units are described in this disclosure to emphasize functional aspects of devices configured to perform the disclosed techniques, but do not necessarily require realization by different hardware units. Rather, as described above, various units may be combined in a codec hardware unit or provided by a collection of interoperative hardware units, including one or more processors as described above, in conjunction with suitable software and/or firmware.

CLAIMS

1. A method of decoding an encoded video sequence, comprising:
obtaining a value of a reference picture list (RPL) syntax element from a bitstream of the video
5 sequence, the RPL syntax element specifying whether syntax elements related to reference
picture lists are present in the bitstream;
when the RPL syntax element specifies that the syntax elements related to reference picture
lists are present in the bitstream: a) obtaining, from the bitstream, values of the syntax
elements related to reference picture lists and b)
10 constructing at least one reference picture list for inter prediction using the obtained values of
the syntax elements related to reference picture lists.
2. The method of claim 1, further comprising:
when the RPL syntax element specifies that the syntax elements related to reference picture
15 lists are not present in the bitstream, constructing at least one reference picture list for inter
prediction without further obtaining values of syntax elements from the bitstream.
3. The method of claim 2, wherein the constructing at least one reference picture list for inter
prediction without further obtaining values of syntax elements from the bitstream
20 comprises performing a sliding window method to construct at least one reference picture
list for inter prediction.
4. The method of any one of claims 1 to 3, further comprising:
when the RPL syntax element specifies that the syntax elements related to reference picture
25 lists are not present in the bitstream, obtaining, from the bitstream, a value of a second syntax
element related to reference picture marking.
5. The method of any one of claims 1 to 4, wherein the syntax elements related to reference
picture lists are first syntax elements related to reference picture lists constructing.
30
6. The method of claim 5, wherein the first syntax elements related to reference picture lists
constructing are included in at least one reference picture list syntax structure, wherein each
reference picture list syntax structure is used to construct one reference picture list.

7. The method of claim 4, wherein the second syntax element related to reference picture marking provides information on a maximum number of reference pictures usable for inter prediction of pictures of the video sequence.

5

8. The method of any of the preceding claims, wherein the RPL syntax element is a flag and a first value of the flag specifies that syntax related to reference picture lists is present in the bitstream and a second value of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.

10

9. The method of claim 8, wherein the first value is 1 and the second value is 0.

10. The method of any of the preceding claims, wherein the value of the RPS syntax element is obtained from one of a sequence parameter set, picture parameter set and slice header.

15

11. The method of any of the preceding claims, wherein the syntax elements related to reference picture lists are obtained from a sequence parameter set and/or a slice header.

12. The method of any of the preceding claims, further comprising, when the RPL syntax element specifies that syntax related to reference picture lists is present in the bitstream, obtaining a value of at least one third syntax element from a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header.

13. The method of claim 12, wherein the at least one third syntax element comprises two third syntax elements, wherein the value of one of the two third syntax elements is obtained for a first reference picture list and, in case that a fourth syntax element obtained from the bitstream indicates that another third syntax element is present in the bitstream, the value of the other one of the two third syntax elements is obtained for a second reference list.

30

14. The method of claim 13, wherein the fourth syntax element is another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

15. The method of claim 12, 13 or 14, wherein the at least one value of the third syntax element is obtained in case that there is at least one reference picture list syntax structure signalled at sequence parameter set level in the bitstream, wherein each reference picture list syntax structure is used to construct one reference picture list.

5

16. A method of encoding a video sequence, comprising:

determining whether values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction;

generating a bitstream comprising a RPL syntax element specifying whether syntax
10 elements related to reference picture lists are present in the bitstream, wherein the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream in case that it is determined that the values of the syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

15

17. The method of claim 16, wherein the bitstream comprises a second syntax element related to reference picture marking, when it is determined that values of syntax elements related to reference picture lists are not to be used for construction of at least one reference picture list for inter prediction.

20

18. The method of claim 16 or 17, wherein the syntax elements related to reference picture lists are first syntax elements related to reference picture lists constructing.

19. The method of claim 18, wherein the first syntax elements related to reference picture lists
25 constructing are included in at least one reference picture list syntax structure, wherein each reference picture list syntax structure is used for construction of one reference picture list.

20. The method of claim 17, wherein the second syntax element related to reference picture marking provides information on a maximum number of reference pictures usable for inter
30 prediction of pictures of the video sequence

21. The method of any one of claims 16 to 20, wherein the RPL syntax element is a flag and a first value of the flag specifies that syntax related to reference picture lists is present in the

bitstream and a second value of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.

22. The method of claim 21, wherein the first value is 1 and the second value is 0.

5

23. The method of any one of claims 16 to 22, wherein the RPL syntax element is signalled in one of a sequence parameter set, picture parameter set and slice header.

24. The method of any one of claims 16 to 23, wherein the syntax elements related to reference picture lists are comprised in a sequence parameter set and/or a slice header.

10

25. The method of any one of claims 16 to 24, wherein at least one third syntax element is comprised in a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header, when it is determined that values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

15

26. The method of claim 25, wherein the bitstream comprises a fourth syntax element and the at least one third syntax element comprises two third syntax elements, wherein one of the two third syntax elements is for a first reference picture list and the other one of the two third syntax elements is for a second reference list in case that a value of the fourth syntax element indicates that the other third syntax element is present in the bitstream.

20

27. The method of claim 26, wherein the fourth syntax element is another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

25

28. The method of any of claims 25 to 27, wherein at least one reference picture list syntax structure is signalled at sequence parameter set level in the bitstream, wherein each reference picture list syntax structure is used for construction of one reference picture list.

30

29. An encoder (20) comprising processing circuitry for carrying out the method according to any of claims 16 to 28.

30. A decoder (30) comprising processing circuitry for carrying out the method according to any of claims 1 to 15.

31. A computer program product comprising program code for performing the method
5 according to any one of the claims 1 to 28 when executed on a computer or a processor.

32. A decoder, comprising:
one or more processors; and
a non-transitory computer-readable storage medium coupled to the processors and storing
10 programming for execution by the processors, wherein the programming, when executed by
the processors, configures the decoder to carry out the method according to any one of the
claims 1 to 15.

33. An encoder, comprising:
15 one or more processors; and
a non-transitory computer-readable storage medium coupled to the processors and storing
programming for execution by the processors, wherein the programming, when executed by
the processors, configures the encoder to carry out the method according to any one of the
claims 16 to 28.

20

34. A non-transitory computer-readable medium carrying a program code which, when
executed by a computer device, causes the computer device to perform the method of any one
of the claims 1 to 28.

25

35. A decoding device (1000) for decoding an encoded video sequence, comprising:
an obtaining unit (1010) configured for obtaining a value of a reference picture list (RPL)
syntax element from a bitstream of the video sequence, the RPL syntax element specifying
whether syntax elements related to reference picture lists are present in the bitstream and,
30 when the RPL syntax element specifies that the syntax elements related to reference picture
lists are present in the bitstream, obtaining, from the bitstream, values of the syntax elements
related to reference picture lists; and
a RPL construction unit (1020) for, when the RPL syntax element specifies that the syntax
elements related to reference picture lists are present in the bitstream, constructing at least one

reference picture list for inter prediction using the obtained values of the syntax elements related to reference picture lists.

36. The decoding device (1000) of claim 35, wherein the RPL construction unit (1020) is
5 configured for, when the RPL syntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, constructing at least one reference picture list for inter prediction without further obtaining values of syntax elements from the bitstream.

37. The decoding device (1000) of claim 36, wherein the constructing at least one reference
10 picture list for inter prediction without further obtaining values of syntax elements from the bitstream comprises performing a sliding window method to construct at least one reference picture list for inter prediction.

38. The decoding device (1000) of any one of claims 35 to 37, wherein the obtaining unit
15 (1010) is configured for, when the RPL syntax element specifies that the syntax elements related to reference picture lists are not present in the bitstream, obtaining, from the bitstream, a value of a second syntax element related to reference picture marking.

39. The decoding device (1000) of any one of claims 35 to 38, wherein the syntax elements
20 related to reference picture lists are first syntax elements related to reference picture lists constructing.

40. The decoding device (1000) of claim 39, wherein the first syntax elements related to
25 reference picture lists constructing are included in at least one reference picture list syntax structure, and wherein the RPL construction unit (1020) is configured for using each reference picture list syntax structure to construct one reference picture list.

41. The decoding device (1000) of claim 38, wherein the second syntax element related to
30 reference picture marking provides information on a maximum number of reference pictures usable for inter prediction of pictures of the video sequence.

42. The decoding device (1000) of any of the claims 35 to 41, wherein the RPL syntax element is a flag and a first value of the flag specifies that syntax related to reference picture lists

is present in the bitstream and a second value of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.

43. The decoding device (1000) of claim 42, wherein the first value is 1 and the second value is 0.

44. The decoding device (1000) of any of claims 35 to 43, wherein the obtaining unit (1010) is configured for obtaining the value of the RPS syntax element from one of a sequence parameter set, picture parameter set and slice header.

10

45. The decoding device (1000) of any of the claims 35 to 44, wherein the obtaining unit (1010) is configured for obtaining the syntax elements related to reference picture lists from a sequence parameter set and/or a slice header.

46. The decoding device (1000) of any of claims 35 to 45, wherein the obtaining unit (1010) is configured for, when the RPL syntax element specifies that syntax related to reference picture lists is present in the bitstream, obtaining a value of at least one third syntax element from a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header.

20

47. The decoding device (1000) of claim 46, wherein the at least one third syntax element comprises two third syntax elements, wherein the obtaining unit (1010) is configured for obtaining the value of one of the two third syntax elements for a first reference picture list and, in case that a fourth syntax element obtained from the bitstream indicates that another third syntax element is present in the bitstream, the value of the other one of the two third syntax elements for a second reference list.

25

48. The decoding device (1000) of claim 47, wherein the fourth syntax element is another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

30

49. The decoding device (1000) of claim 46, 47 or 48, wherein the obtaining unit (1010) is configured for obtaining the at least one value of the third syntax element in case that there is at least one reference picture list syntax structure signalled at sequence parameter set level in

the bitstream, and wherein the RPL construction unit (1020) is configured for using each reference picture list syntax structure to construct one reference picture list.

50. An encoding device (1100) for encoding a video sequence, comprising:

5 a determining unit (1110) for determining whether values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction;

10 a bitstream generation unit (1120) configured for generating a bitstream comprising a RPL syntax element specifying whether syntax elements related to reference picture lists are present in the bitstream, wherein the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream in case that it is determined by the determining unit (1110) that the values of the syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

15 51. The encoding device (1100) of claim 50, wherein the bitstream comprises a second syntax element related to reference picture marking, when it is determined by the determining unit (1110) that values of syntax elements related to reference picture lists are not to be used for construction of at least one reference picture list for inter prediction.

20 52. The encoding device (1100) of claim 50 or 51, wherein the syntax elements related to reference picture lists are first syntax elements related to reference picture lists constructing.

25 53. The encoding device (1100) of claim 52, wherein the first syntax elements related to reference picture lists constructing are included in at least one reference picture list syntax structure, wherein each reference picture list syntax structure is used for construction of one reference picture list.

30 54. The encoding device (1100) of claim 51, wherein the second syntax element related to reference picture marking provides information on a maximum number of reference pictures usable for inter prediction of pictures of the video sequence

55. The encoding device (1100) of any one of claims 50 to 54, wherein the RPL syntax element is a flag and a first value of the flag specifies that syntax related to reference picture lists is

present in the bitstream and a second value of the flag different from the first value specifies that syntax related to reference picture lists is not present in the bitstream.

56. The encoding device (1100) of claim 55, wherein the first value is 1 and the second value is 0.

57. The encoding device (1100) of any one of claims 50 to 56, wherein the RPL syntax element is comprised in one of a sequence parameter set, picture parameter set and slice header.

58. The encoding device (1100) of any one of claims 50 to 57, wherein the syntax elements related to reference picture lists are comprised in a sequence parameter set and/or a slice header.

59. The encoding device (1100) of any one of claims 50 to 58, wherein at least one third syntax element is comprised in a slice header specifying whether the syntax elements related to reference picture lists are obtainable from a sequence parameter set or a slice header, when it is determined by the determining unit (1110) that values of syntax elements related to reference picture lists are to be used for construction of at least one reference picture list for inter prediction.

60. The encoding device (1100) of claim 59, wherein the bitstream comprises a fourth syntax element and the at least one third syntax element comprises two third syntax elements, wherein one of the two third syntax elements is for a first reference picture list and the other one of the two third syntax elements is for a second reference list in case that a value of the fourth syntax element indicates that the other third syntax element is present in the bitstream.

61. The encoding device (1100) of claim 60, wherein the fourth syntax element is another flag, a value of 1 of the other flag indicating that the other third syntax element is present in the bitstream.

62. The encoding device (1100) of any of claims 59 to 61, wherein at least one reference picture list syntax structure is comprised in a sequence parameter set, wherein each reference picture list syntax structure is used for construction of one reference picture list.

63. A non-transitory storage medium which includes a coded bitstream, the bitstream being generated by dividing a current picture of a video signal or an image signal a plurality of blocks, wherein the bitstream comprises a plurality of syntax elements, wherein the plurality of syntax elements comprises a RPL syntax element specifying whether syntax elements related to
- 5 reference picture lists are present in the bitstream, the plurality of syntax elements further comprises the syntax elements related to reference picture lists when the RPL syntax element specifies that the syntax elements related to reference picture lists are present in the bitstream.